MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

AD-A142 340

AD-A142 340
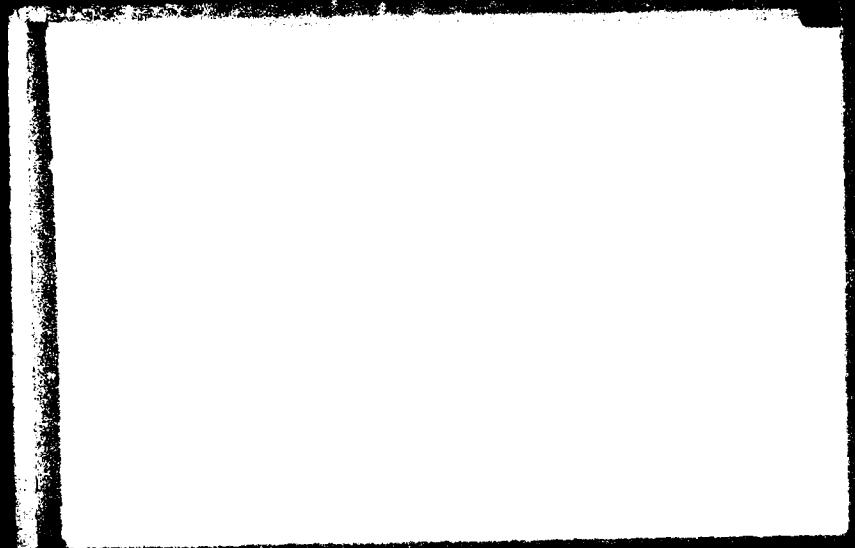
OPERATOR PERFORMANCE ON FLAT-PANEL DISPLAYS

WITH LINE AND CELL FAILURES

Sandra R. Abramson, M. S.
and
Harry L. Snyder, Ph.D.

January, 1984

Department of Industrial Engineering and Operations Research
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

JUL 2 5 1984

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. *AD-A142 346*	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Operator Performance on Flat-Panel Displays With Line and Cell Failures		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER HFL-83-3
7. AUTHOR(s) Sandra R. Abranson, M.S. Harry L. Snyder, Ph.D.		8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0238
9. PERFORMING ORGANIZATION NAME AND ADDRESS Office of Naval Research, Code 455 800 North Quincy Street Arlington, Virginia 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 196-155
11. CONTROLLING OFFICE NAME AND ADDRESS Virginia Polytechnic Institute and State Univ. Human Factors Lab Blacksburg, Virginia 24061		12. REPORT DATE January, 1984
		13. NUMBER OF PAGES 45 + vi
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

visual displays	human engineering	visual performance
flat panel displays	image quality	display reliability
human factors	display design	

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This technical report describes the results of an experiment to determine the influence of individual line and cell failures on readability of flat-panel displays. A 1024 X 1024 pixel plasma display was used to present reading passages. Selected percentages of dots or cells were failed (either on or off) randomly within the passage. Reading time for these passages was measured and related to the type of failure, percent failure, character font, and case of the characters (upper and mixed).

-- The results clearly indicate that all these variables have an effect upon readability of the display, but that interactions among the variables are complex and important. Design criteria for displays which are subject to such failures are offered, as are quality assurance criteria.

PREFACE

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

The effectiveness of the visual presentation of information depends upon several categories of variables, among them the symbolic representation of the information, the typography, and the information content. Numerous experiments and analyses have been conducted over the years to define the relationships between legibility and readability of displays and such design variables as character and symbol size, strokewidth, contrast, font or character style, height/width ratio, and numerous others. McCormick and Sanders (1982) have summarized many of these results, while entire volumes have been written on the subject of character font alone (e.g., Cornog and Rose, 1967). As a result of these experimental data, we have a reasonably good understanding of how to design characters and symbols for maximum legibility on a printed page.

However, with the advent of electronic displays, new constraints exist upon the placement and composition of characters and symbols, constraints which cannot be addressed by existing data on printed materials. Among these constraints is the fact that most electronic displays create characters from a series of either lines or dots, rather than with continuous strokes as is the case with printed text. An additional constraint is the fact that, for many electronic displays, the width or thickness of the dot or stroke is fixed for the entire display and the designer's problem is to choose how many discrete widths to use for a single character "strokewidth," rather than to choose a continuously variable width of the single strokewidth.

Thus, the legibility of electronic display symbols is affected by variables quite different from those of printed text. These influences of these differences on character and text legibility have not been adequately addressed in the literature, with the result that commercially available electronic displays have no standards for character design and presentation. Nor, unfortunately, does one find any consistency in such designs, either among manufacturers or within the product line of a single manufacturer.

The general problem described above is exacerbated by a fundamental difference between printed and electronic displays--namely, the occasional tendency for electronic displays to fail locally. That is, some electronic displays will fail by having certain portions or elements of the display remain in the "on" or "off" state irrespective of the intended state of that display location. Thus, in spite of good design of the individual characters, such failures may in fact contribute to reduced legibility or readability. As the failures increase in number, the display becomes logically less legible and therefore less usable. Unfortunately, data to support acceptability decisions and product quality assurance are unavailable, such that the user or purchaser is left with a decision to accept or reject a partially failed display with no supporting quantitative basis or data.

1

The research described in this report addresses this problem to a limited extent. Specifically, it addresses the effects of certain types of failures of electronic display elements on the readability of text. Because importantly related variables are also investigated in the research, the folowing discussion summarizes current knowledge of the influence of these several variables on text readability.

Character Font

Character font or style can have a significant effect on both legibility and readability of characters and symbols. As pointed out by Cornog and Rose (1967), numerous visually attractive fonts are poorly desigred for legibility and readability. Conversely, the most readable fonts are frequently those which seem most uninteresting.

Electronic displays limit the font designer to a great extent. Some electronic displays, as noted above, provide a finite strokewidth, which virtually eliminates the use of serifs, curved lines, and other discriminating characteristics. Flat panel displays limit font design to an even greater extent than do some cathode-ray tube devices. In particular, flat panel displays create alphanumeric characters and symbols by illuminating selected picture elements (pixels) or dots in a defined cell matrix. Thus, a typical design might contain seven horizontal dots and nine vertical dots in a matrix which can be used to define any given character or symbol. The font must be selected such that **all** characters in the alphabet, all numerals, and all symbols of interest are formed from combinations of dots within this 7 X 9 rectilinear matrix. Other displays use a 5 X 7 matrix, while some use even larger matrices, such as 9 X 11 or 13 X 17.

Maddox, Burnette, and Gutmann (1977) compared three different 5 X 7 dot matrix constrained fonts to find an optimum font for upper case alphanumerics. They found that the maximum dot font resulted in fewer confusions than did the maximum angle font or the Lincoln/Mitre font. Because the maximum dot font used, as the name implies, more dots to create each character than did the other fonts, they concluded that the improved legibility was due to the greater perceived (spatially integrated) contrast of this font. Snyder and Maddox (1978), in a subsequent and similar study, evaluated the Lincoln/Mitre, maximum dot, maximum angle, and Huddleston fonts in three different matrix sizes, 5 X 7, 7 X 9, and 9 X 11. Controlling differently for exposure time, they found that the 7 X 9 and 9 X 11 matrix sizes were more legible than was the 5 X 7 size. For the 5 X 7 size, the Huddleston font was most legible, while the Lincoln/Mitre and Huddleston were superior to the others for both the 7 X 9 and 9 X 11 matrix sizes.

Both Shurtleff (1970) and Snyder and Maddox (1978) have demonstrated the superiority of the Lincoln/Mitre font. The Huddleston (1970) font was developed for high ambient illumination environments by combining the most legible characters from each of several fonts that were experimentally evaluated. Unfortunately, literature does not exist on the legibility of a number of fonts used on commercially available electronic displays, although anecdotal information

2

suggests that many suffer from poor character design and reduced legibility.

Related directly to the legibility of dot matrix characters are numerous other display variables, such as dot size, inter-dot spacing, dot shape, and character physical size. For experimental data on these and other parameters, the reader is referred to Snyder and Maddox (1978) and Snyder (1980).

Matrix Size

Vartabedian (1973) noted that a matrix larger than 5 X 7 is needed to provide sufficient dots for legible lower case characters. Since text printed with both upper case and lower case characters is read faster than is text printed with only upper case characters (Tinker, 1965; Vartabedian, 1970), it is desirable to define optimum electronic display dot matrix fonts for both upper and lower case characters. Thus, the evaluation of matrix sizes for reading text should investigate the utility of upper and lower case fonts.

Conversely, single words and characters are located faster in a search display if they are designed as all upper case characters (Vartabedian, 1970). As pointed out by Albert (1975), studies of single character legibility or single word legibility do not necessarily translate their results into full text readability. Thus, font optimization for text readability may produce results different than font optimization for single character or single word legibility (Scanlan and Carel, 1976), as the manner by which words and sentences are constructed and the relative frequency of words have considerable influence on text readability and word discrimination.

Flat Panel Display Failures

The nature of flat panel displays varies among the technologies used to construct the display. Similarly, the propensity for failure for each technology differs. For example, AC plasma displays tend to fail by having discrete cells remain "off" regardless of their intended state. Similarly, thin-film transister addressed electro-luminescent displays can fail in a single cell or pixel mode and also in a complete line, either vertical or horizontal. Matrix addressed displays tend to fail a line at a time, either vertical or horizontal, depending on the failed driver location and display orientation. Thus, there exists the potential (and demonstrated existence) of individual cell failures, horizontal line failures, and vertical line failures.

Many flat panel displays are capable of both positive (light characters on a dark background) and negative (dark characters on a light background) contrast. In the case of failed cells or lines that cannot be turned "on", the detectability of the failure depends upon whether the cell should be "on" or "off," which may depend also on the positive or negative contrast selection made. Thus, under certain conditions, failed cells may be nondetectable while under other circumstances failed cells may be quite noticeable.

There are three obvious (Pastor and Uphaus, 1982) outcomes of display cell or line failures: (1) the user will correctly identify the character, but perhaps require a longer time to do so; (2) the user will be unable to identify or read the character or word; and (3) the user will confuse the degraded character or word with another character or word, thereby misreading it. Both Bauer (1962) and Long, Reid, and Queal (1951) demonstrated that randomly added or removed dots reduce character identification. Riley and Barbato (1978) reported that only small differences in single character legibility were found among the different 5 X 7 degraded dot matrix characters that they investigated. Their fonts included the ASCII, Lincoln/Mitre, Huddleston, Ellis, and NAMEL. They also found that correct identification of characters was not differentially affected by the addition or removal of dots, or by simultaneously adding and removing dots. Pastor and Uphaus (1982) found a linear relationship between specific dot loss and reading failures. As background noise levels composed of random dots increased in four different typefaces, Spencer, Reynolds, and Coe (1977) demonstrated that readability of passages also decreased.

As suggested by the above summary, the effects of cell failure have been partially investigated, using only upper case characters and certain fonts. No data are known to exist on the individual or combined effects of line failures, upper and lower case characters, and positive vs. negative contrast.

Readability Measurement

This experiment used a modification of the Tinker Speed of Reading Test, a test which had been used with considerable success in previous research (Burnette, 1976; Snyder and Maddox, 1978). The Tinker test is an accurate and reliable measure of operator reading performance with electronic displays for the following reasons: (1) the baseline time per passage can be subtracted from overall reading time to eliminate initial between-subjects differences; (2) subjects are usually familiar with reading tasks such as this, so there are minimal learning effects; and (3) the Tinker test can be used in a variety of display formats and applications. As indicated by Tinker (1965), the test is of uniform difficulty and is not confounded by problems of passage comprehension.

To achieve a valid and generalizable measure of the influence of the display variables on readability, the passages were presented in the same location on the display each time. The failed cells did not always occur in the region of the incongrous word in the passage, simply because we did not wish to artificially constrain the locations of line and cell failures. That is, on an operational failed display, the failure locations would appear randomly and independent of the location of information text presentation; thus, a similar unconstrained location selection was used in this experiment. Cell or line failures in parts of the operational display other than the part containing the passage might distract the subject so that the resulting reading performance might be degraded, although the

4

legibility of the specific text would be unaffected. Similarly, we did not constrain the failed locations for that reason.

Thus, the experiment as conducted was designed to investigate the effects of various failure types and forms on text readability without artificial display location constraints. For that reason, the experiment provides data useful in estimating readability reductions caused by these variables and criteria for acceptance of displays for continued operational usage.

Objectives of the Experiment

The objectives of this experiment were to determine the effects of cell and line failures on readability of flat panel displays with particular emphasis on the following display variables:

1. Upper case vs. mixed (upper and lower) case text
2. Character font
3. Failure type (cell, horizontal line, vertical line)
4. Percent cells failed
5. Failure mode (on, off)

METHOD

Subjects

Sixty college students (30 male) attending Virginia Polytechnic Institute and State University participated as subjects. These volunteers were paid for their participation. All subjects were screened for corrected 20/20 near-point visual acuity using a Bausch and Lomb Orthorater.

Reading Task

A Tinker Speed of Reading Test (Carver and Tinker, 1970) was used to measure the readability of the displayed text. Each subject was given 96 different reading passages. All passages were made up of one or two sentences of 30 words, including a word that does not make any sense and does not fit into the context of the sentence. Appendix A contains some examples of these passages.

Procedure

Subjects were screened for acceptable visual acuity prior to their taking part in the experiment. During the experiment, each subject sat in front of an AC plasma panel display and was given a short set of written instructions (Appendix B). Seven practice trials with all the experimental variables were given to allow time for the subject to adjust to the darkened room illumination level. All practice trials were given in the same manner as in the main experiment. The practice session required about five minutes, during and following which questions could be asked.

A triggering device was located adjacent to the plasma display. The trigger was depressed and released to cause one passage to appear on the display. After reading the passage and locating the inappropriate word, the subject again pressed the trigger and verbally identified the inappropriate word. With the second press of the response trigger the displayed passage was erased. A new passage was then available for display upon the subject's depressing the trigger once again.

Each passage was displayed singly with a particular combination of the experimental variables. The subject had no time limit per passage. All 96 passages were randomized so that no two subjects had the same order of presentation. The experimental combinations were also randomized to prevent the occurrence of the same combinations on two consecutive trials.

The experimenter remained in the room throughout the entire experiment and was seated in front of an HP2621A terminal to monitor the subject's responses. Correct, incorrect, or no (null) identification of the inappropriate word in the passage was recorded on the terminal for each trial. The computer automatically recorded the response time, the

time between the subject's first release of the trigger and his second depression of it. To avoid or reduce boredom, subjects received two 15-20 minute experimental sessions with a rest period between the sessions.

Experimental Design

Three within-subjects variables--failure mode, percent cells failed, and failure type--were investigated. Failure mode had two levels, on and off. An on failure resulted in a cell being inappropriately illuminated, while an off failure resulted in a cell being inappropriately not illuminated. Percent failure consisted of six levels--0, 4, 8, 12, 16, and 20 percent of the cells being failed either on or off. Failure type was controlled at three levels--discrete cell failures, horizontal line failures, and vertical line failures.

Two between-subjects variables, font and case, were also investigated. The three fonts were the Lincoln/Mitre, Huddleston, and HP. The HP font is the font found on the HP2621A computer terminal. Case consisted of either upper case only or a mized upper and lower case in which only the first letter of the first word of a sentence was capitalized.

The locations of cell and line failures were randomly selected on each trial. Thus, failed cells or lines did not always occur in the region of the inappropriate word that the subject had to identify. Failed cells and lines were randomly programmed separately for each subject and for each experimental condition.

Each subject was assigned to one of the three fonts and to either the mixed case or upper case condition. Within these six combinations, subjects received each combination of the within-subjects variables (3 failure types x 2 failure modes x 5 percent failures + 1 zero failure condition) a total of 3 times. Hence, each subject received 96 experimental trials, as illustrated in Figures 1 and 2.

In each cell of the experimental design data were collected for three trials. The mean of these three scores was used in subsequent analyses, so that the total number of scores in the statistical analyses was 2160.

PERCENT FAILURE
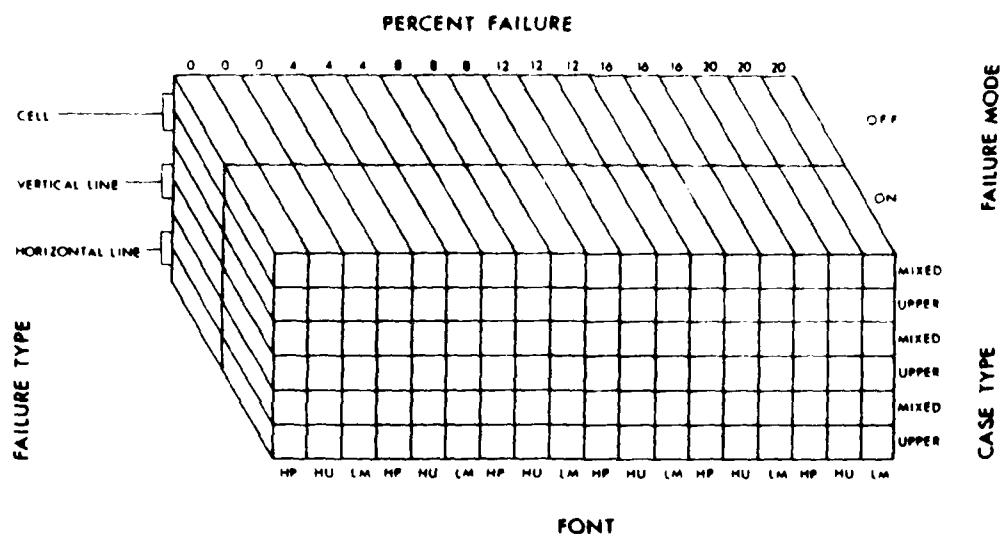
Figure 1. Overall Experimental Design.
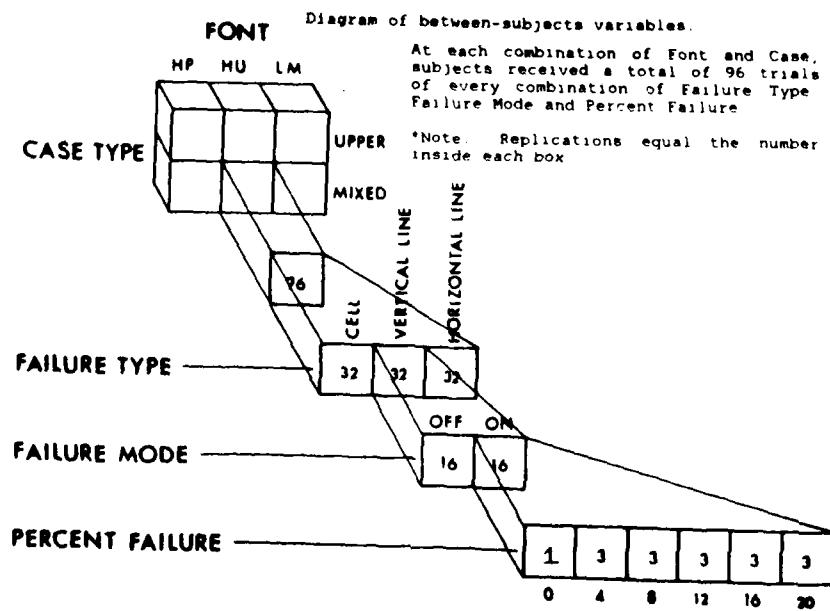
Figure 2. Experimental Design, Within-Subjects Breakdown.

Creation of Lower Case Fonts

No documentation has been found on lower case Huddleston or Lincoln/Mitre fonts. For maximum legibility, it is necessary to optimize both upper and lower case characters. Since larger character sets will result in greater likelihoods of between-character confusion (Snyder, 1980), an optimal upper case font may not remain optimal when a mixed case font is substituted for the upper case characters. Thus, care must be taken in selecting characters that are both consistent with font appearance as well as with a minimum-confusion criterion. Because the intended purpose of this experiment was not simply font optimization, no attempt was made to optimize either the Huddleston or Lincoln/Mitre lower case font. (The HP font was used in its current form, optimized or not.) Thus, lower case characters were generated to maintain the internal consistency of each of these fonts and to preserve characteristics similar to the upper case characters. Lower case characters were written with two dot descenders.

The HP2621A font was modified for its use on the plasma panel. Symbols generated on the plasma display are restricted in that they have equally spaced vertical and horizontal pixel locations. By comparison, the HP2621A terminal has horizontal spacing smaller than its vertical spacing. The HP characters were therefore modified to have the same aspect ratio as on the HP terminal, but using fewer horizontal dots to conform to the geometry of the plasma display.

Equipment

A PDP 11/10 minicomputer and custom software were used to drive the plasma display. It was placed on a standard height (75 cm) laboratory table, with no direct room illumination falling on the display.

Plasma display. The Photonics Technology AC plasma panel had 1024 by 1024 pixels, with a resolution of 24 pixels per centimeter both vertically and horizontally. The total display active area is therefore 43.2 cm square. Using the 7 X 9 dot matrix character cell, it is possible to write 32 lines of 64 characters per line. The maximum dot luminance is about 206 cd/m^2, with a contrast ratio of at least 25:1. The display emission is neon orange in color, with a dominant wavelength of 585.2 nm. When illuminated, adjacent dots cannot be visually resolved at normal viewing distance due to their irregular shape and nonuniform intensity distribution (Snyder, 1980).

All passages and failure conditions appeared in a centered display window of 368 pixels wide and 200 pixels high, as illustrated in Figure 3. The passages were written in either five or six 30-character lines. No words were hyphenated.

RESULTS

Response Time

An analysis of variance was performed on the response times, followed by simple-effects F-tests for significant interactions and post hoc Newman-Keuls tests on significant main effects and simple-effect results (Tables 1 through 6).

TABLE 1. Analysis of Variance Summary for Response Times

Source of Variance	df	MS	F	p
Between Subjects				
Font (F)	2	336.7	1.2	0.3053
Case (C)	1	2413.1	8.7	0.0047
F x C	2	157.5	0.6	0.5704
Subjects within Font,				
Case (S/F,C)	54	277.6		
Within Subjects				
Failure Type (FT)	2	1077.4	97.3	0.0001
Percent Failure (PF)	5	2110.8	143.2	0.0001
Failure Mode (FM)	1	2119.6	114.9	0.0001
FT x F	4	32.0	2.9	0.0256
PF x F	10	29.6	2.0	0.0327
FM x F	2	2.2	0.1	0.8887
FT x C	2	209.6	18.9	0.0001
PF x C	5	186.6	10.1	0.0001
FM x C	1	301.0	16.3	0.0002
FT x PF	10	247.9	25.3	0.0001
FT x FM	2	2799.5	143.3	0.0001
FM x PF	5	290.9	28.7	0.0001
FT x S/F,C	108	11.1		
PF x S/F,C	270	14.7		
FM x S/F,C	54	18.4		
FT x PF x FM	10	571.2	51.7	0.0001
FT x PF x F	20	8.7	0.9	0.6111
FT x FM x F	4	55.2	2.9	0.0266
FT x PF x C	10	36.1	3.7	0.0001
FT x FM x C	2	423.5	22.0	0.0001
FM x PF x F	10	15.7	1.6	0.1214
FM x PF x C	5	24.7	2.4	0.0351

TABLE 1 (continued)

Source of Variance	df	MS	F	p
FT x PF x S/F,C	540	9.8		
FT x FM x S/F,C	108	19.3		
FM x PF x S/F,C	270	10.1		
F x C x FM	2	11.8	0.6	0.5311
F x C x PF	10	26.4	1.8	0.0620
F x C x FT	4	17.1	1.6	0.1937
FT x PF x FM x F	20	19.5	1.8	0.0218
FT x PF x FM x C	10	67.2	6.1	0.0001
FT x PF x FM x S/F,C	540	11.1		
F x C x FM x PF	10	8.1	0.8	0.6314
F x C x FT x FM	4	27.1	1.4	0.2366
F x C x FT x PF	20	10.0	1.0	0.4327
FT x PF x FM x F x C	20	9.3	0.8	0.6599
Total	2159			

Cell failures resulted in the longest response times, while horizontal line failures led to the shortest response times (Figure 3). All differences among failure type means are statistically significant (p < .01).

Although the font main effect is not significant, several interactions involving font are significant. The font by failure type interaction is illustrated in Figure 4, which indicates that the Huddleston is better than the other two fonts for both cell failures and vertical line failures, but that no statistically significant differences exist (Table 3) for horizontal line failures. Differences between the HP and Lincoln/Mitre fonts are not significant for any of the failure types.

As the percent of failed cells increases, the response time consistently increases (Figure 5). The greater the percent failure the longer is the response time for all fonts, except at 0% and 4% for the Huddleston font, as illustrated in Figure 6.

RESPONSE TIME, SECONDS
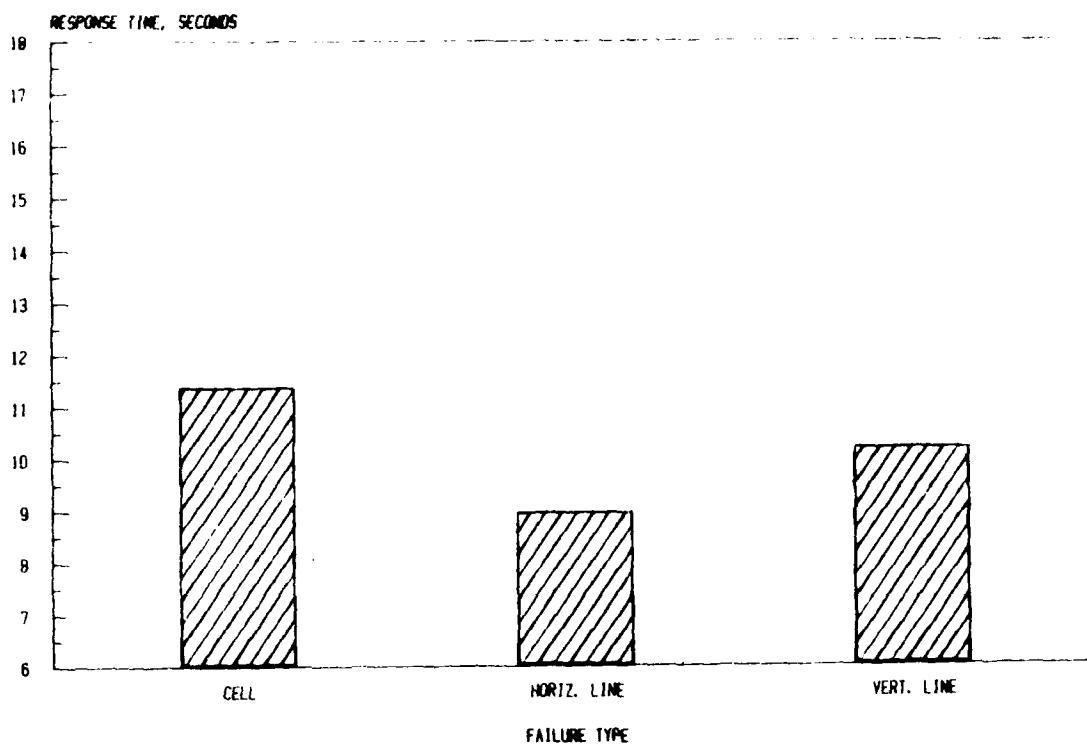
Figure 3.   Effect of Failure Type on Response Time.
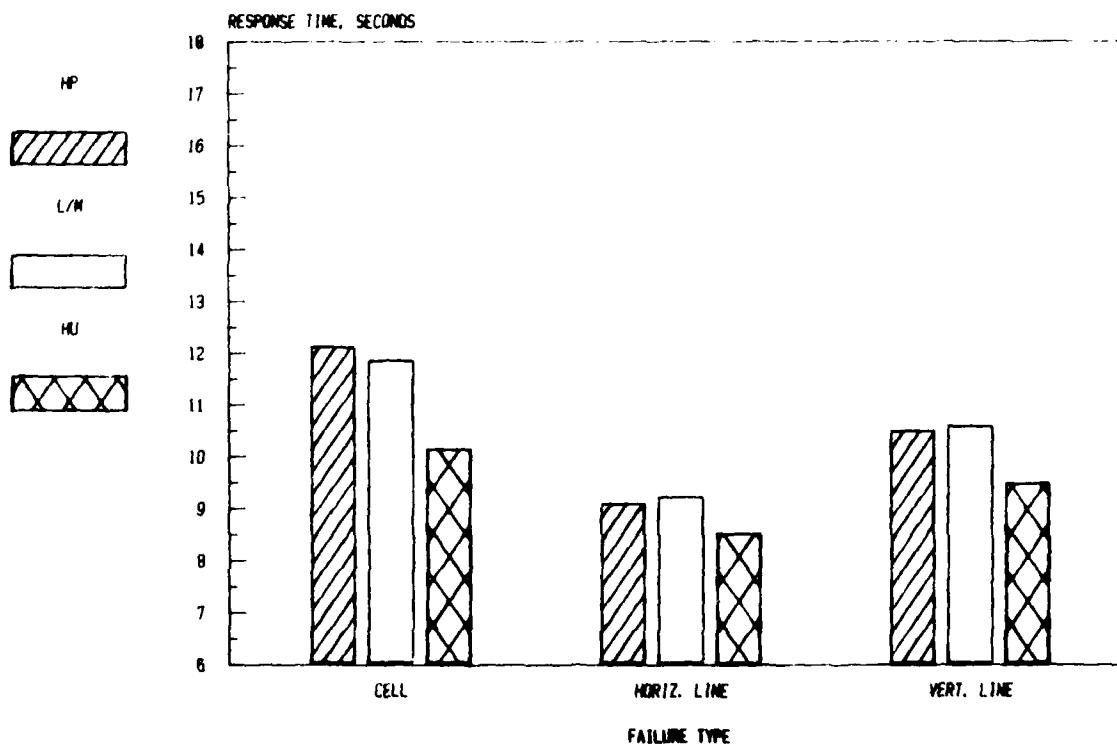

Figure 4.   Effect of Font and Failure Type on Response Time.
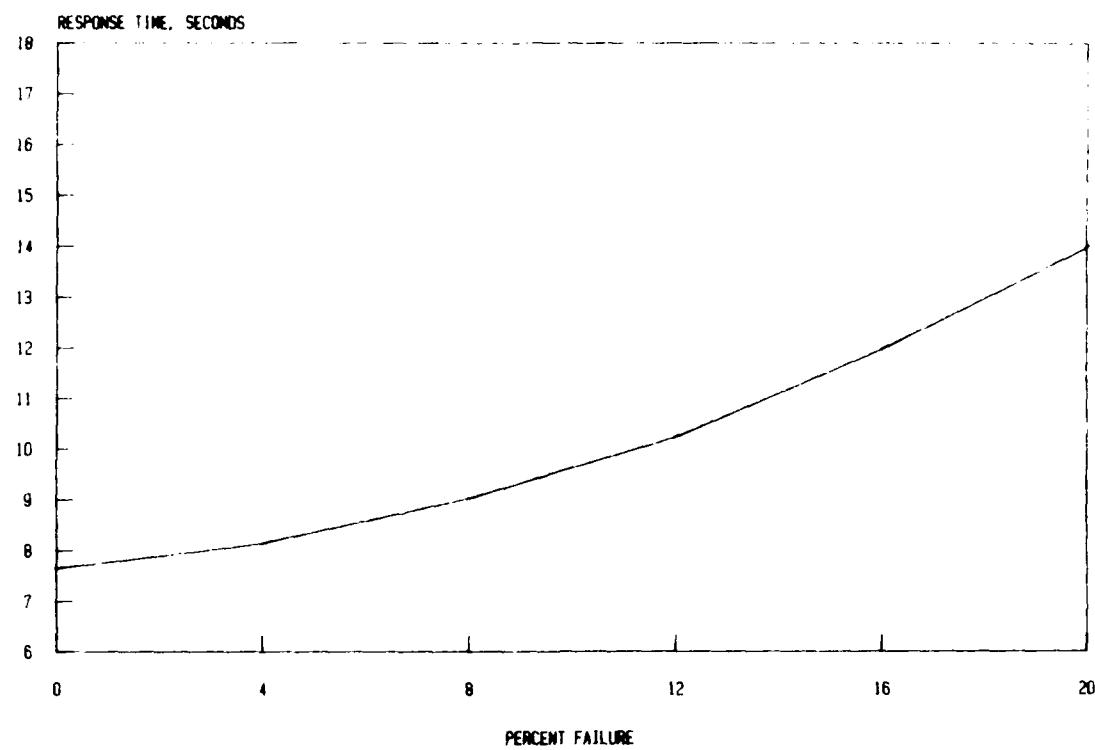
Figure 5. Effect of Percent Failures on Response Time.
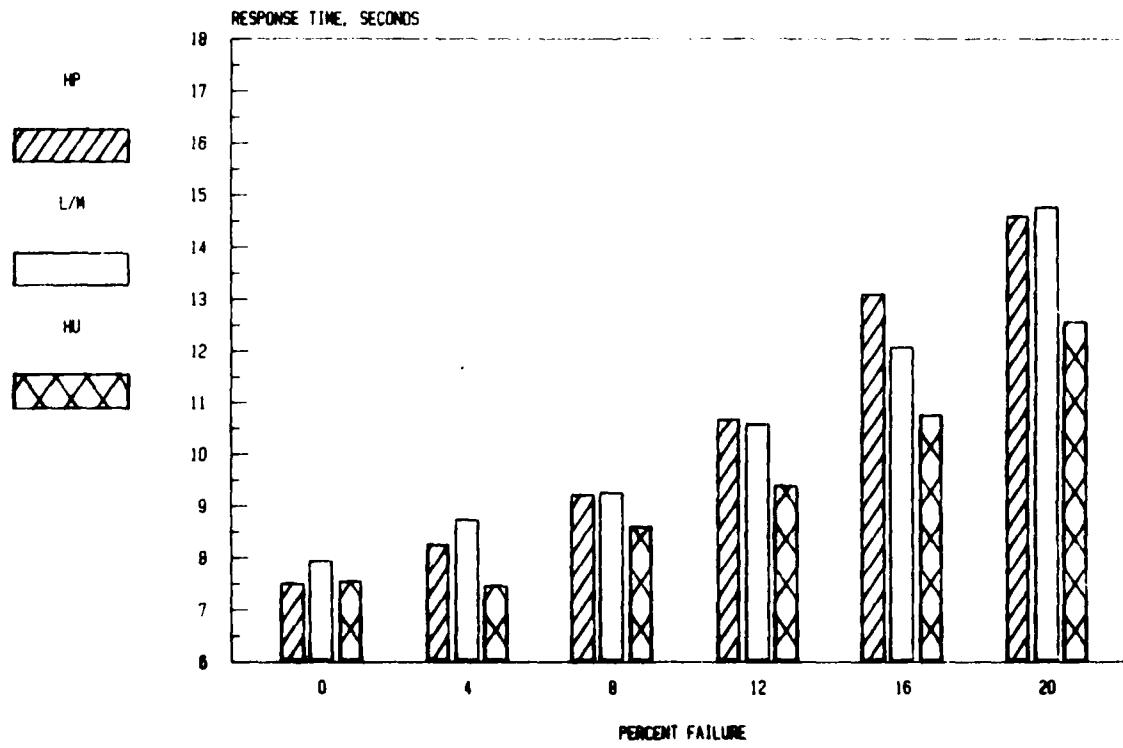
Figure 6. Effect of Font and Percent Failures on Response Time.

13

The Huddleston font consistently has the shortest response time for all percent failure levels. There is a significant difference between the Lincoln Mitre and the Huddleston fonts for 4%, 12%, 16%, and 20% (p < .05), and a significant difference between the Lincoln Mitre and the HP fonts for the 16% failure level (p < .01). As illustrated in Figure 6, increases in percent cells failed leads to increases in differences among the fonts.

Upper case passages were read in less time than were mixed case passages (Figure 7). While this difference is consistent for all three failure types, the magnitude of the difference is greatest for the cell failures and smallest for the horizontal line failures, as shown in Figure 8. Response times are significantly shorter for the horizontal line failures than for the other failure types for both cases (p < .01), and the difference between the vertical and horizontal line failures is significant (p < .01) for the mixed case, but not for the upper case (p > .05).

As the percent cells failed increases, the response times increase more for the mixed case than for the upper case characters for failure percents 4 through 20 (p < .01, Figure 9). In addition, response times are significantly greater with the on cell failures than with the off cell failures for both upper and mixed cases (p < .01, Figure 10), although the magnitude of the effect is greater for the on failures.

While response times are consistently longer for mixed case characters, the pattern is greatly influenced by combinations of failure type and percent cells failed, as illustrated in Figure 11. Longer response times are associated with larger proportions of failed cells, this trend being greater for mixed case passages. However, the differences among the failure percents and among the failure types are reduced considerably for upper case passages, such that selection of an all upper case font compensates for between 4% and 8% of additionally failed cells.

Response times were considerably longer for on cell failures for all combinations of failure type and case (Figure 12). However, "on" failures led to the best performance with horizontal line failures and the poorest with cell failures, while "off" failures produced the best performance with cell failures and the poorest with vertical line failures. Performance was poorer with both cell and vertical line "on" failures than with any of the "off" failure combinations of case and failure type.

As the percent failed cells increased, the advantages of upper case presentation over mixed case also increased for both on and off failure modes, as illustrated in Figure 13. The upper case response times were generally about 20 to 30 percent less than were comparable mixed case response times for all percent failure and failure mode combinations.

14
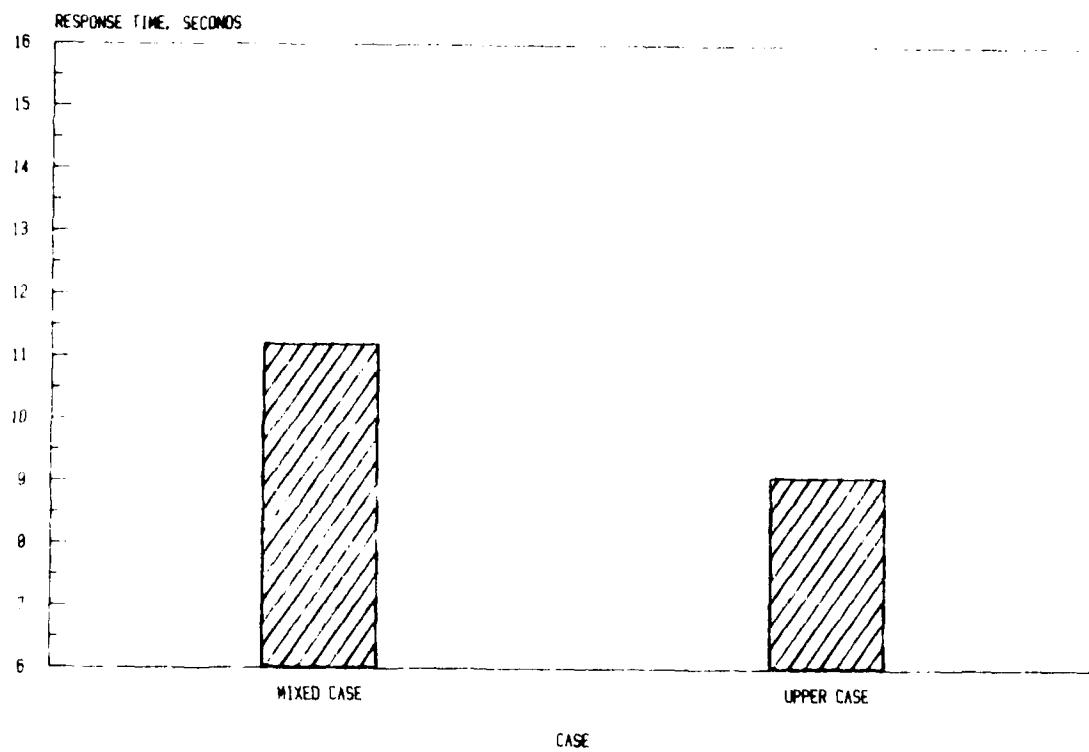
Figure 7. Effect of Case on Response Time.
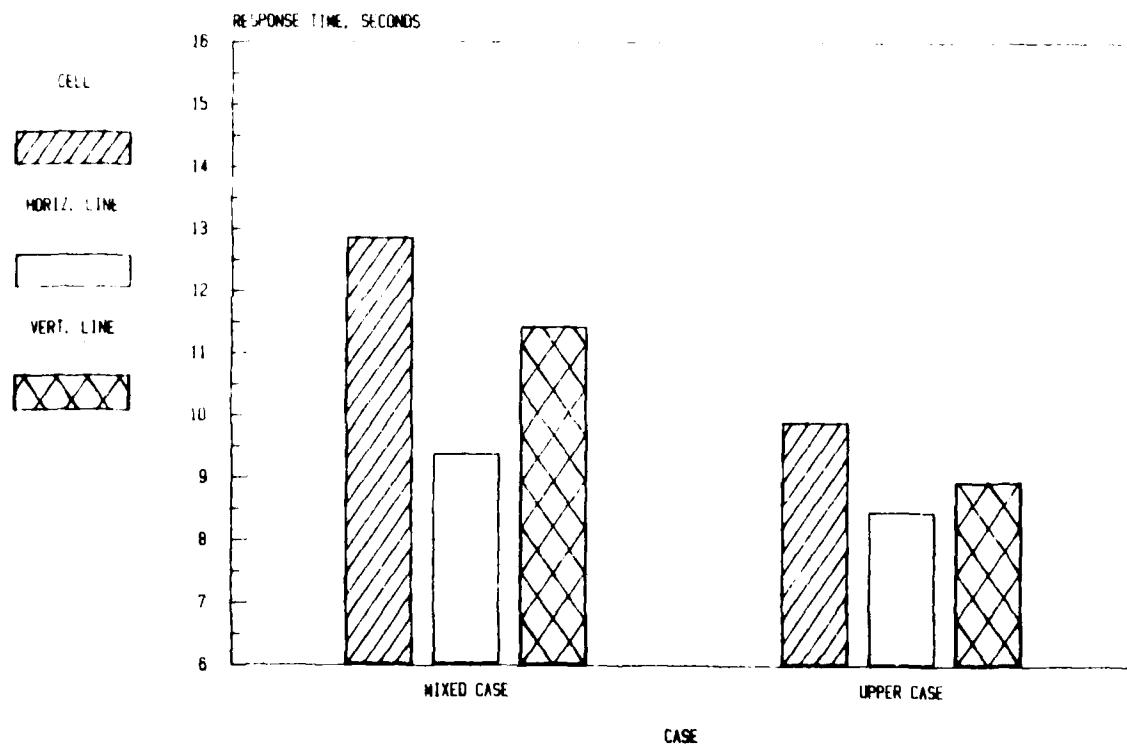
Figure 8. Effect of Case and Failure Type on Response Time.

15

RESPONSE TIME, SECONDS

MIXED CASE

UPPER CASE
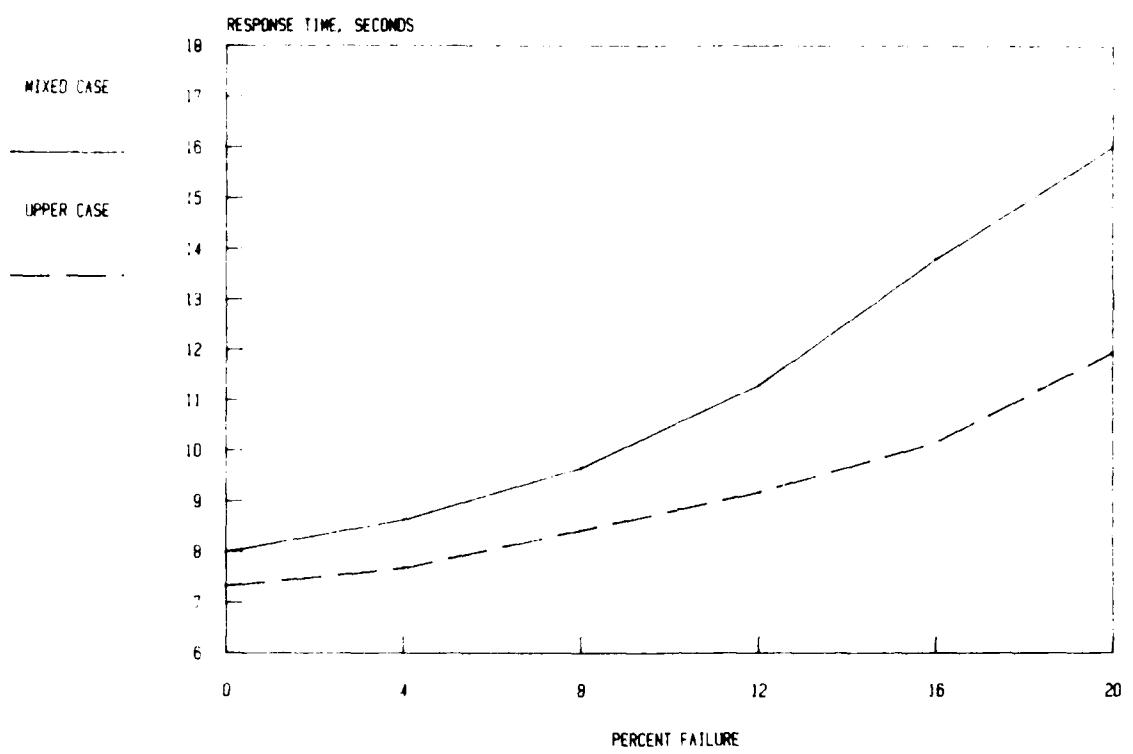
PERCENT FAILURE

Figure 9.   Effect of Case and Percent Failure on Response Time.

RESPONSE TIME, SECONDS

MIXED CASE   UPPER CASE
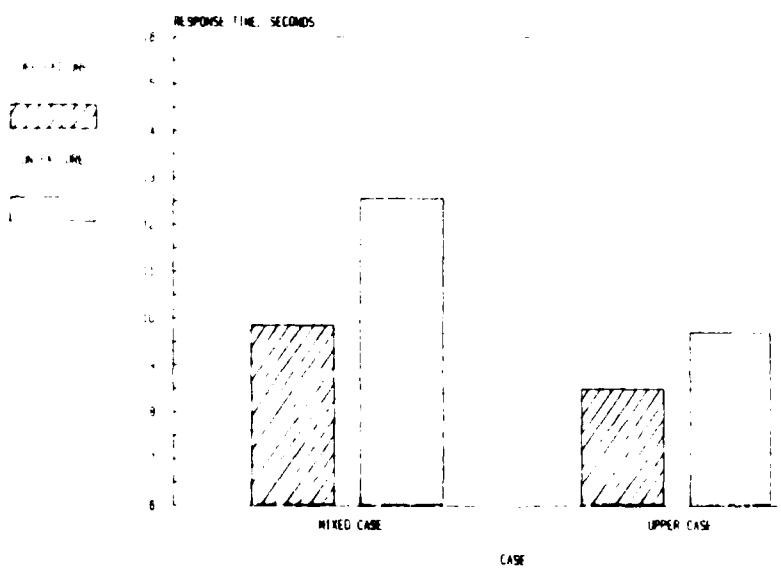
CASE

Figure 10.   Effect of Case and Failure Mode on Response Time.

16
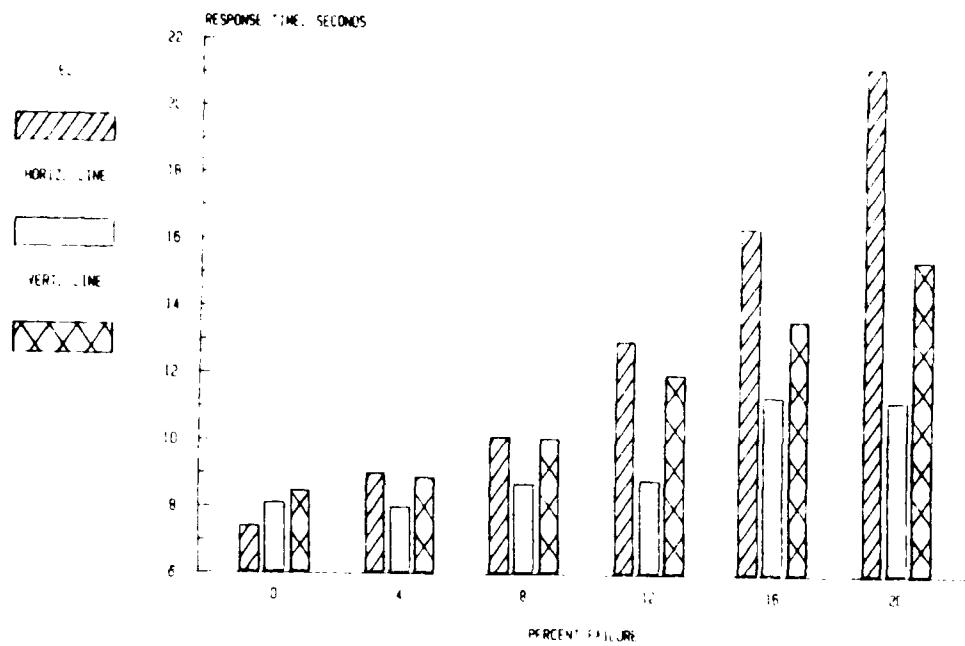
Figure 11a.   Effect of Failure Type and Percent Failure on
Response Time for Mixed Case Text.
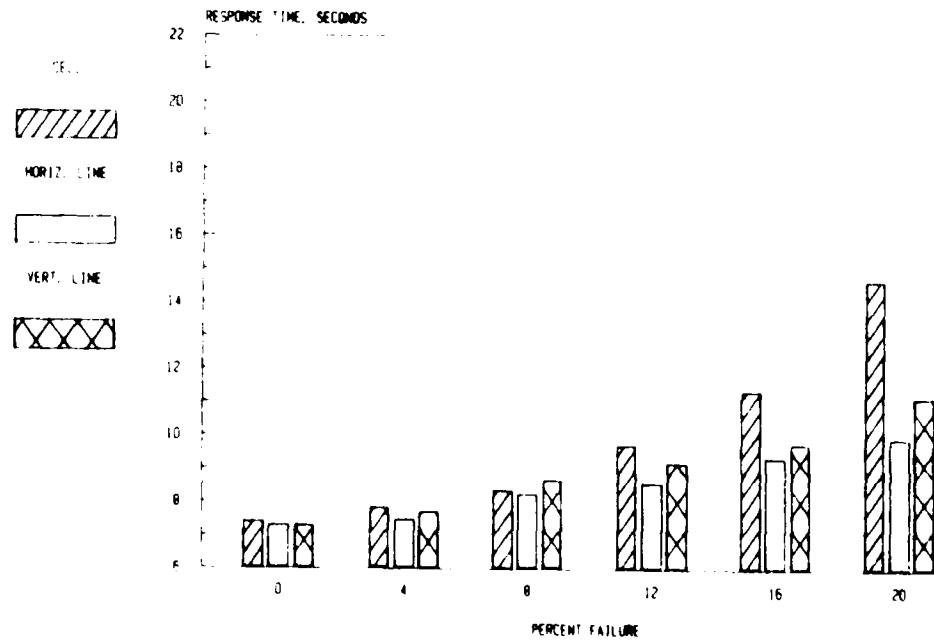
Figure 11b.   Effect of Failure type and Percent Failure
on Response Time for Upper Case Text.

17
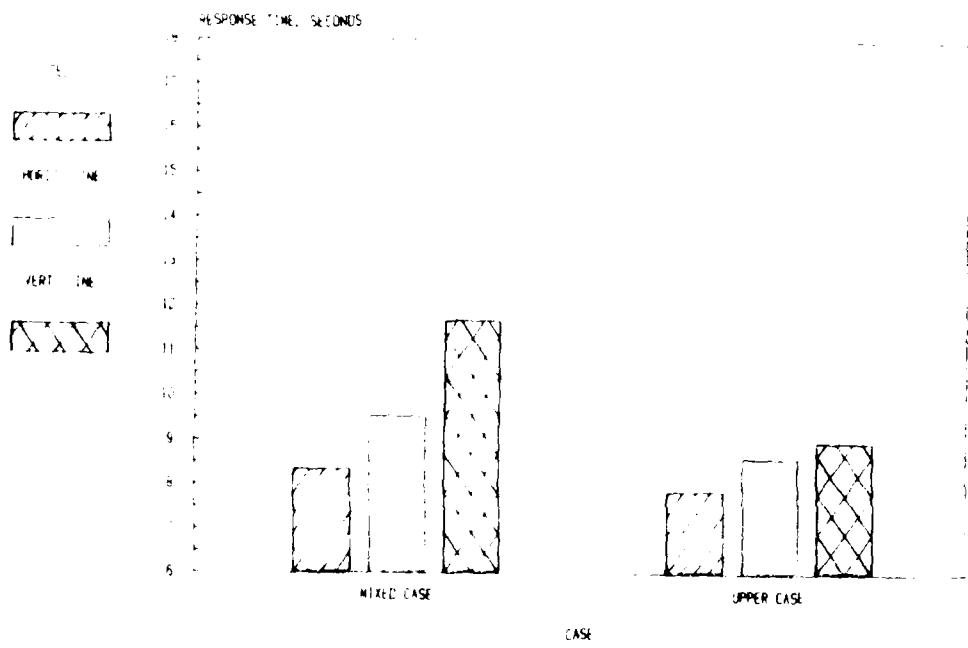
Figure 12a. Effect of Case and Failure Type on Response
Time for the Off Failure Mode.
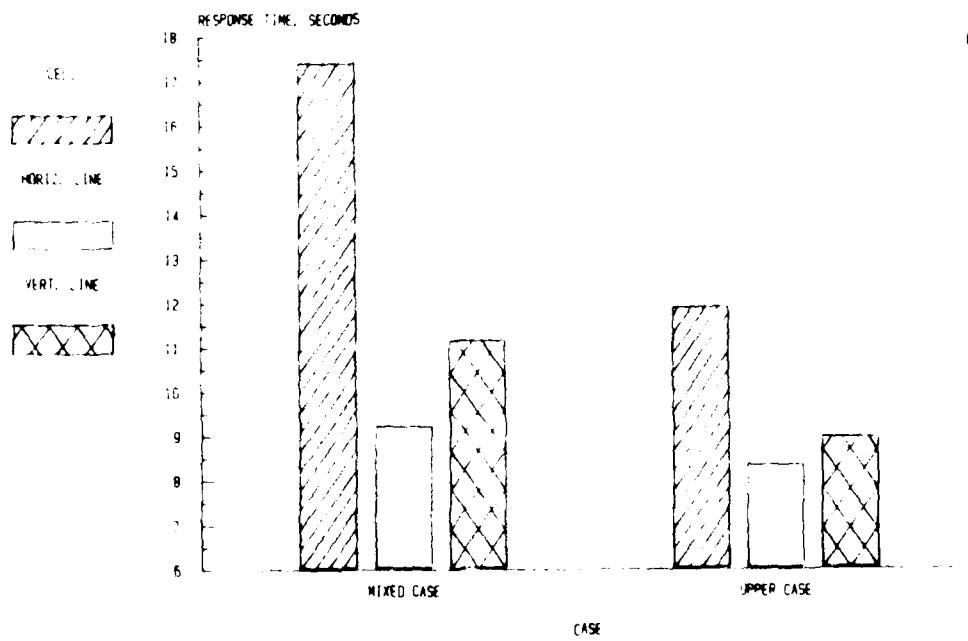
Figure 12b. Effect of Cast and Failure Type on Response
Time for the On Failure Mode.

18
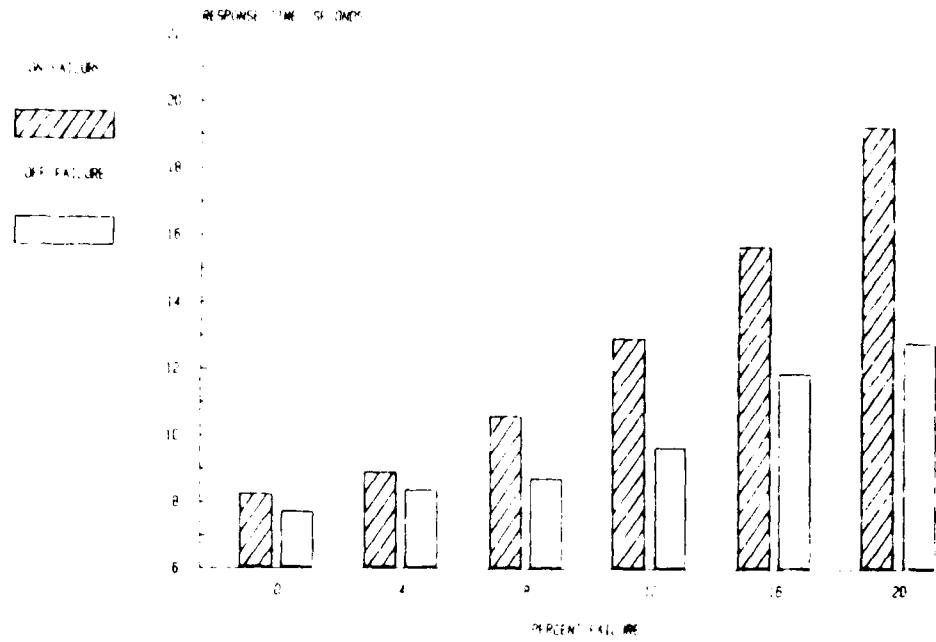
Figure 13a.   Effect of Failure Mode and Percent Failure on
Response Time for Mixed Case Text.
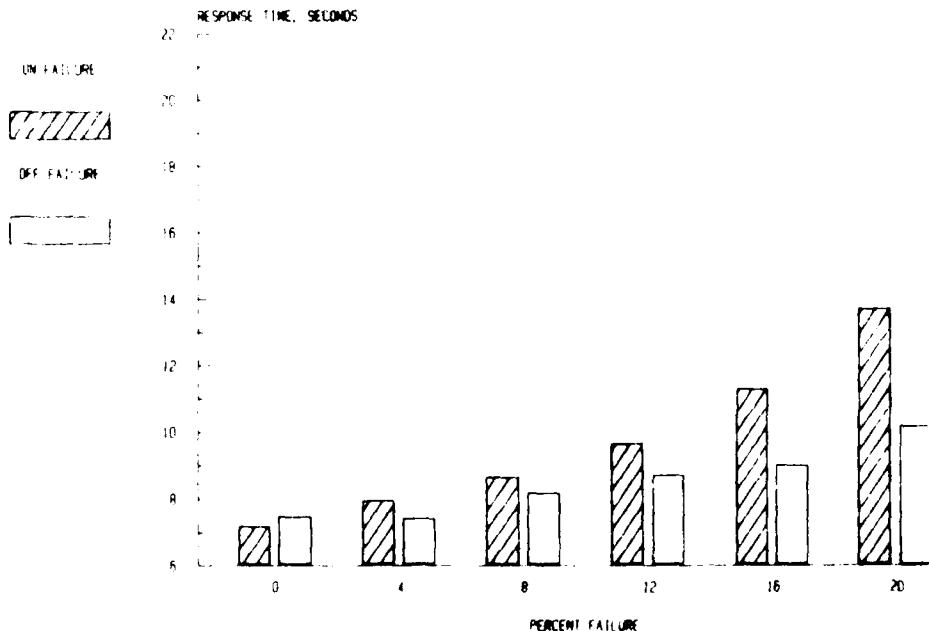
Figure 13b.   Effect of Failure Mode and Percent Failure on
Response Time for Upper Case Text.

19

Although the font main effect is not statistically significant, the shortest response times occurred consistently with the Huddleston font for all combinations of failure type and failure mode, as illustrated in Figure 14. The superiority of the Huddleston font is greater as the overall response times increase, for example with "on" cell failures or "off" vertical line failures.

"Off" failures in general result in shorter response times than do "on" failures (Figure 15), but this effect is very dependent on the failure type. As shown in Figure 16, cell failures result in the shortest response times for "off" cells, but in the longest response times for "on" failures. The effect of failure mode is not significant for either vertical or horizontal line failures. The differences between vertical and horizontal line failure response times are significant (p < .01) for both on and off failure modes, while the difference between cell failures and horizontal line failures is not significant for the "off" failure mode (p > .05).

"On" failures result in increasingly greater response times than do "off" failures as the number of such failures increases (Figure 17). The failure mode effect is small and not statistically significant (p > .05) for failure rates of 4% or less, but becomes much larger as failure percent increases to the point where a 20% failed cell percentage results in a 35% increase in response time for on failures compared to off failures.

The variables investigated in this experiment necessarily interact significantly with one another in their effects on response time, such that main effect conclusions are of limited generality. Nowhere is this more apparent than in the examination of the interaction among failure type, percent cells failed, and failure mode, as illustrated in Figure 18. As indicated, cell failures have very little effect on response times for off cell failures. However, for on cell failures, there is a very great effect of the percent cells failed. In fact, the effect of percent cells failed is small for horizontal line on and off failures, and somewhat larger for off vertical line failures than for on vertical line failures. The relationship between this pattern and character structure will be discussed in a later section of this report.
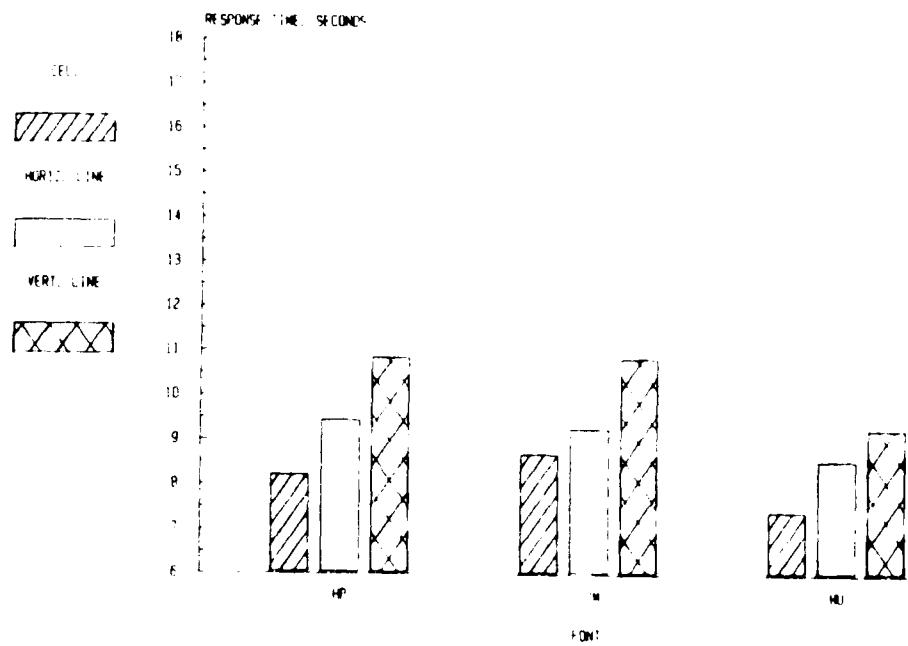
Figure 14a.   Effect of Font and Failure Type on Response
Time for the Off Failure Mode.
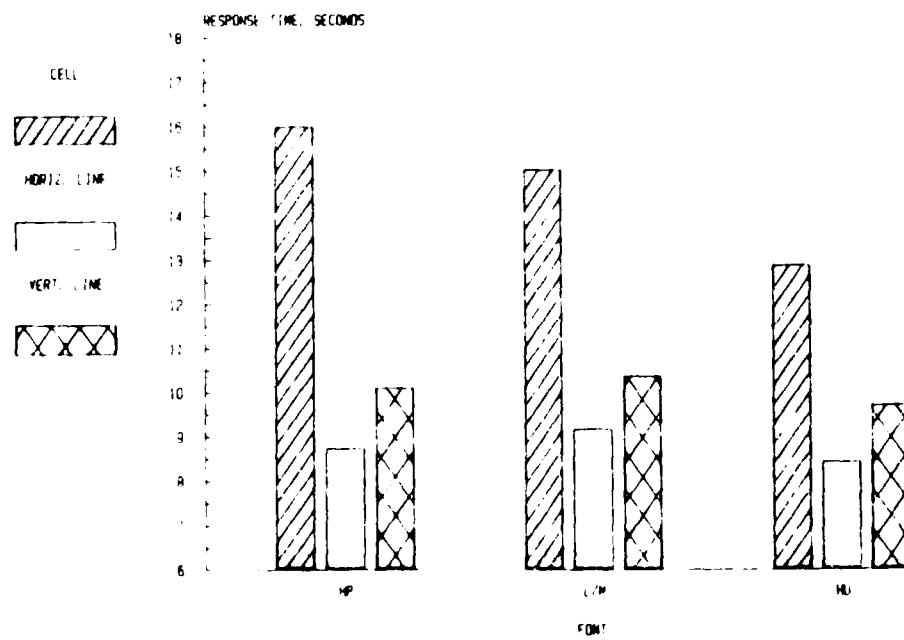
Figure 14b.   Effect of Font and Failure Type on Response
Time for the On Failure Mode.

21
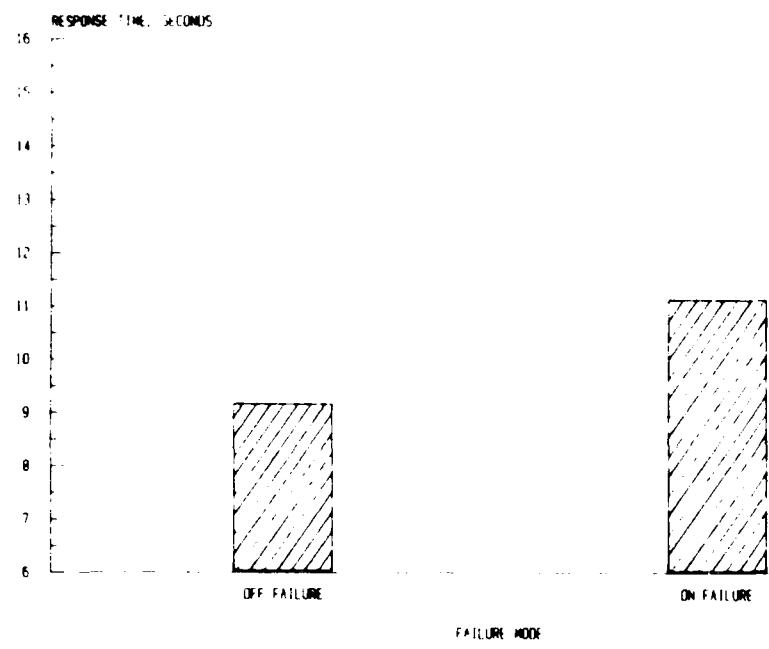
Figure 15.   Effect of Failure Mode on Response Time.
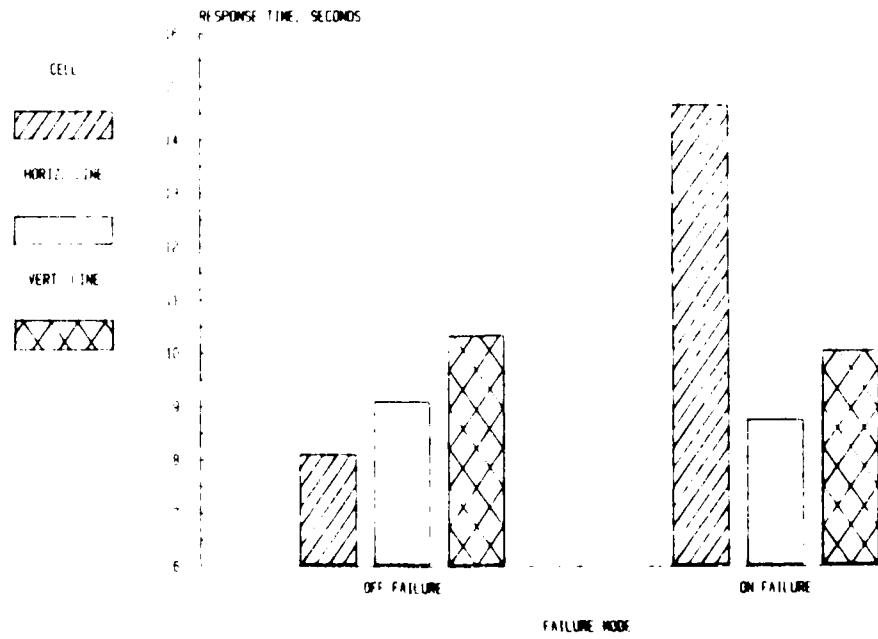
Figure 16.   Effect of Failure Mode and Failure Type
on Response Time.
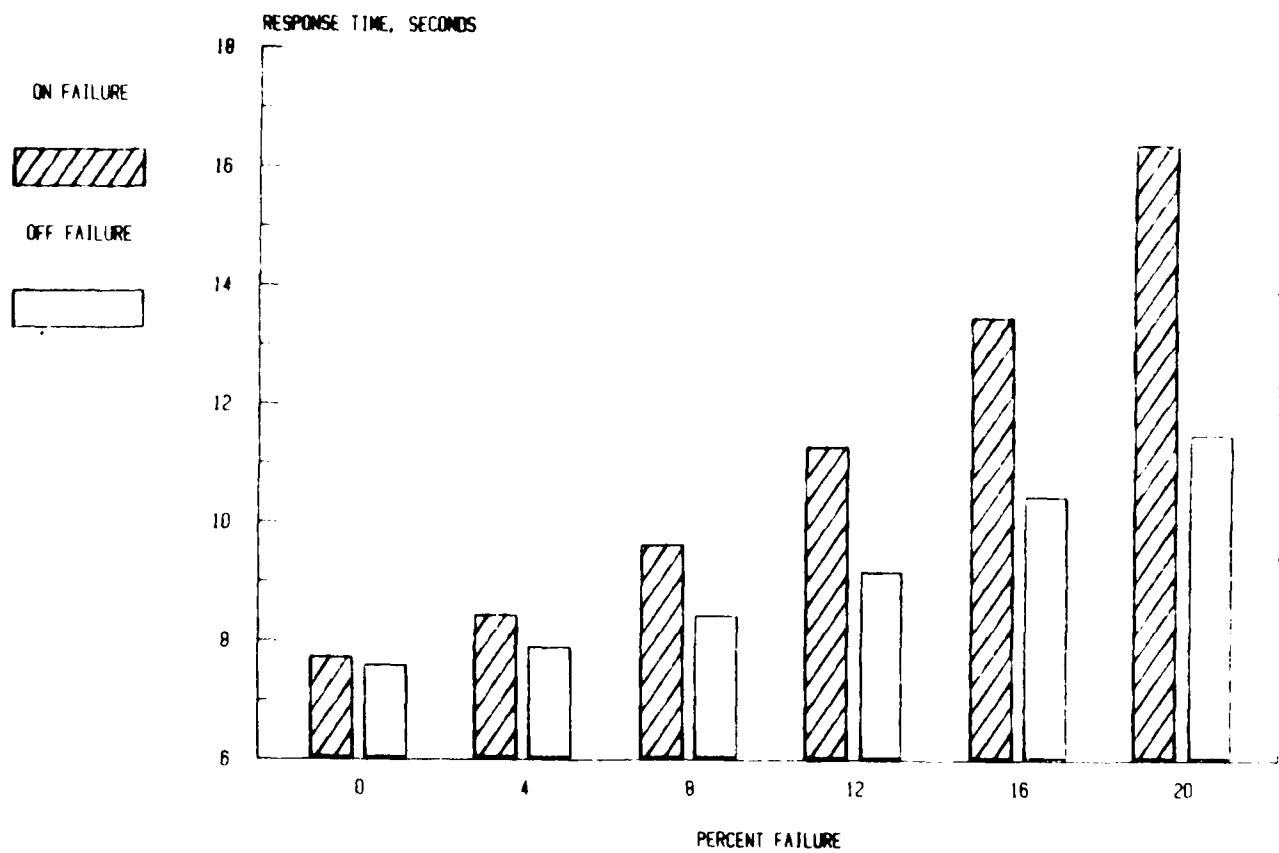
Figure 17. Effect of Failure Mode and Percent Failure
on Response Time.
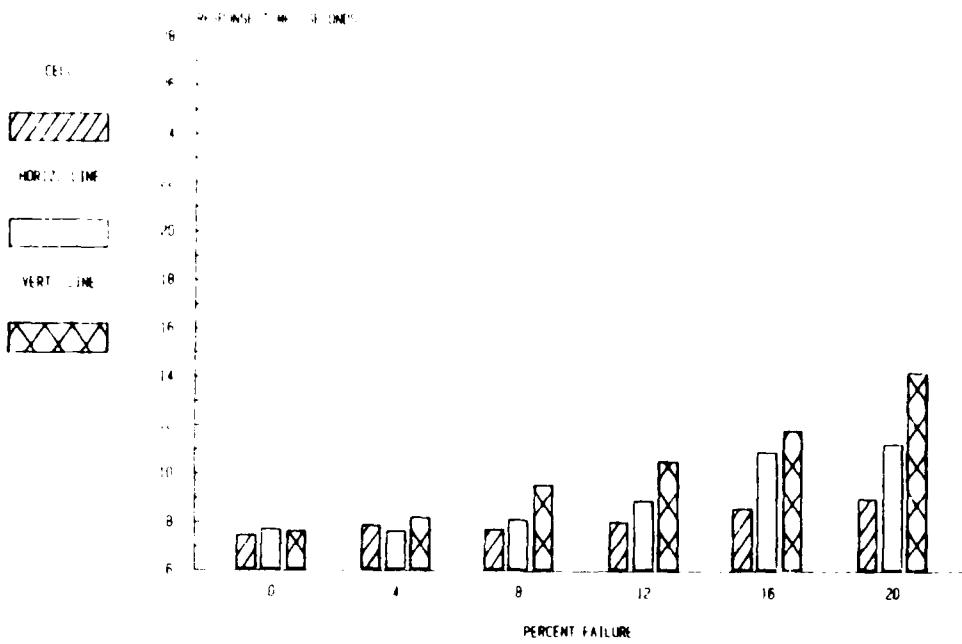
Figure 18a. Effect of Failure Type and Percent Failure on
Response Time for the Off Failure Mode.
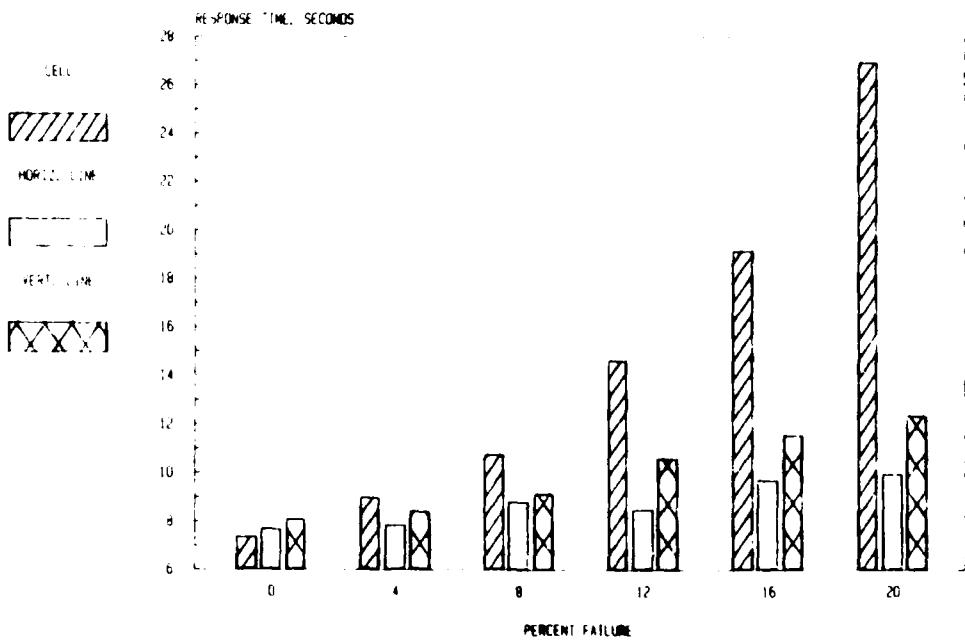
Figure 18b. Effect of Failure Type and Percent Failures on
Response Time for the On Failure Mode.

24

TABLE 2. Summary Table for the Individual Simple-Effect
F-Tests for Each Font

Source of Variance	df	MS	F
Font = HP			
FT	2	556.1	50.2*
PF	5	932.6	63.2*
FM x FT	2	1450.9	75.3*
Font = Huddleston			
FT	2	165.9	15.0*
PF	5	468.1	31.7*
FM x FT	2	567.6	29.5*
Font = Lincoln/Mitre			
FT	2	419.5	37.9*
PF	5	769.3	52.2*
FM x FT	2	891.5	46.3*

* $p < .001$

TABLE 3. Summary Table for the Individual Simple-Effect
           F-Tests for Each Failure Type

Source	df	MS	F
Cell Failures			
F	2	273.9	24.7*
C	1	1590.6	143.7*
PF	5	1889.0	128.1*
FM	1	7687.1	416.9*
F x FM	2	82.9	4.3
C x PF	5	186.2	19.0*
C x FM	1	1136.2	59.0*
FM x PF	5	696.4	63.0*
Horizontal Line Failure			
F	2	35.0	3.2
C	1	146.7	13.3*
PF	5	191.4	13.0*
FM	1	17.5	0.9
F x FM	2	7.4	0.4
C x PF	5	13.0	1.3
C x FM	1	0.0	0.0
FM x PF	5	19.7	1.8
Vertical Line Failure			
F	2	91.8	8.3*
C	1	1095.0	98.9*
PF	5	526.2	35.7*
FM	1	14.0	0.8
F x FM	2	22.4	1.2
C x PF	5	59.8	6.1*
C x FM	1	11.6	0.6
FM x PF	5	20.9	1.9

* $p < .01$

TABLE 4. Summary Table for the Individual Simple-Effect
F-Tests for Each Percent Failure Level

Source	df	MS	F
0% Failure			
F	2	6.6	0.4
C	1	40.0	2.7
FT	2	6.6	0.7
FM	1	1.1	0.1
C x FT	2	11.5	1.2
C x FM	1	15.7	1.5
FM x FT	2	2.4	0.2
4% Failure			
F	2	48.1	3.3
C	1	82.4	5.6
FT	2	17.1	1.7
FM	1	24.7	2.4
C x FT	2	4.0	0.4
C x FM	1	0.0	0.0
FM x FT	2	7.2	0.7
8% Failure			
F	2	17.1	1.2
C	1	135.5	9.2*
FT	2	26.8	2.7
FM	1	123.5	12.2*
C x FT	2	14.4	1.5
C x FM	1	41.5	4.1
FM x FT	2	92.2	8.3*

* $p < .01$

27

TABLE 4, continued.

Source	df	MS	F
12% Failure			
F	2	64.1	4.3
C	1	405.0	27.5*
FT	2	220.5	22.5*
FM	1	396.5	39.1*
C x FT	2	79.6	7.7*
C x FM	1	120.6	11.8*
FM x FT	2	468.1	42.3*
16% Failure			
F	2	164.6	11.2*
C	1	1194.5	81.0*
FT	2	375.6	38.3*
FM	1	833.3	82.2*
C x FT	2	69.0	7.0*
C x FM	1	52.0	5.1
FM x FT	2	1274.2	115.3*
20% Failure			
F	2	184.3	12.5*
C	1	1489.9	101.0*
FT	2	1670.5	170.1*
FM	1	2195.3	216.4*
C x FT	2	211.6	21.5*
C x FM	1	194.4	19.2*
FM x FT	2	3811.4	344.8*

* p < .01

28

TABLE 5.   Summary Table for the Individual Simple-Effect
           F-Tests for Each Case

Source	df	MS	F
Upper Case			
FT	2	185.5	16.7*
PF	5	531.5	36.0*
FM	1	411.6	22.3*
FT x PF	10	53.6	5.5*
FM x FT	2	526.2	27.3*
FM x PF	5	86.4	8.5*
Mixed Case			
FT	2	1101.5	99.5*
PF	5	1766.2	119.8*
FM	1	2009.0	108.9*
FT x PF	10	230.4	23.5*
FM x FT	2	2696.7	140.0*
FM x PF	5	229.2	22.6*

* $p < .01$

TABLE 6. Summary Table for the Individual Simple-Effect
          F-Tests for Each Failure Mode

Source	df	MS	F
"On" Failures			
C	1	2209.3	119.8*
FT	2	3431.1	178.1*
PF	5	1982.9	195.4*
F x FT	4	77.5	3.9*
C x FT	2	504.7	26.2*
C x PF	5	156.9	15.5*
FT x PF	10	759.2	68.7*
"Off" Failures			
C	1	504.8	27.4*
FT	2	445.8	23.1*
PF	5	419.7	41.4*
F x FT	4	9.8	0.5
C x FT	2	128.4	6.7*
C x PF	5	54.6	5.4*
FT x PF	10	59.9	5.4*

* $p < .01$

Response Frequency

Responses to the inappropriate word in the reading passages were categorized as correct, incorrect, or null. An incorrect response was a false positive, a response of a word other than the word that was correct. A null response occurred when the subject was unable to read the inappropriate word in the passage. Separate chi-square tests were conducted for incorrect responses (Table 7) and null responses (Table 8) for each of the five independent variables. Factorial chi-square analyses were not performed because of the small expected frequencies per cell. While analyses were performed on the numbers of responses, figures are plotted in terms of percent of total responses to provide more information to the reader.

The proportion of on failure responses exceeds that of off responses, as shown in Figure 19. However, the difference between on and off incorrect responses is not significant ($p > .05$). Most of the incorrect and null responses were made with cell failures and the fewest were always made with horizontal line failures (Figures 20 and 21).

As the percent of failed cells increases, so does the proportion of incorrect responses, as shown in Figure 22. The proportion of null responses increases with percent failed cells in excess of 8%, but levels below 8% do not affect the null response rate (Figure 23).

As with response times, both incorrect response and null response performance is better with upper case than with mixed case passages, as illustrated in Figures 24 and 25.

It should be noted that these significant differences in incorrect and null responses represent less than 1% of the total responses. Thus, while the differences attributable to the experimental variables are statistically significant, most of the influence of the experimental variables on performance is on response time, not on the likelihood of a correct response.

31

TABLE 7.   Chi-Square Summary Table for Incorrect Responses

Source	Chi-Square	df	p
Failure Mode	3.56	1	> .05
Failure Type	11.08	2	< .01
Percent Failure	34.00	5	< .001
Case	6.72	1	< .01
Font	3.58	2	> .10

TABLE 8.   Chi-Square Summary Table for Null Responses

Source	Chi-Square	df	p
Failure Mode	17.29	1	< .001
Failure Type	41.52	2	< .001
Percent Failure	42.33	5	< .001
Case	19.44	1	< .001
Font	2.38	2	> .40

PERCENT NULL RESPONSES
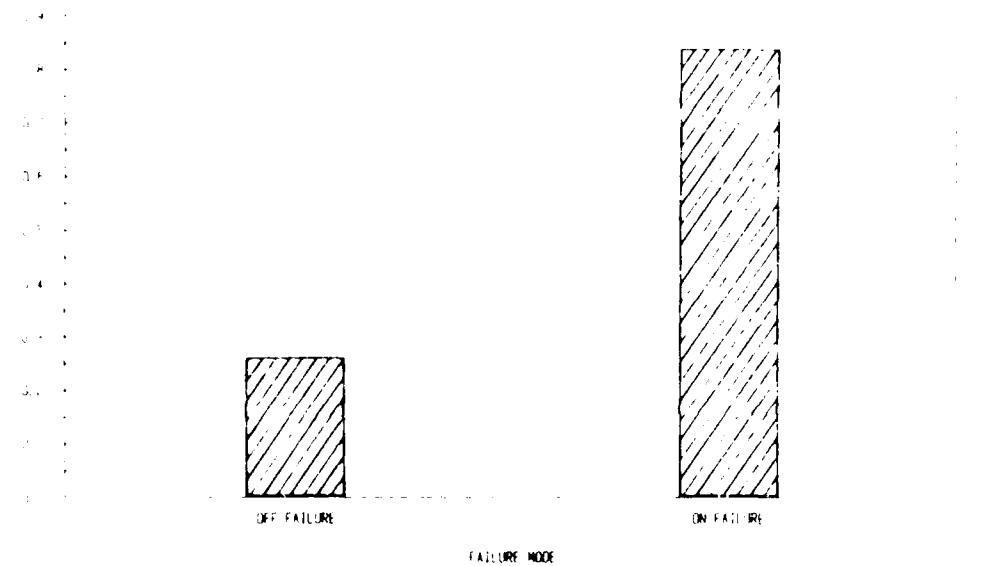
Figure 19. Effect of Failure Mode on Null Responses.
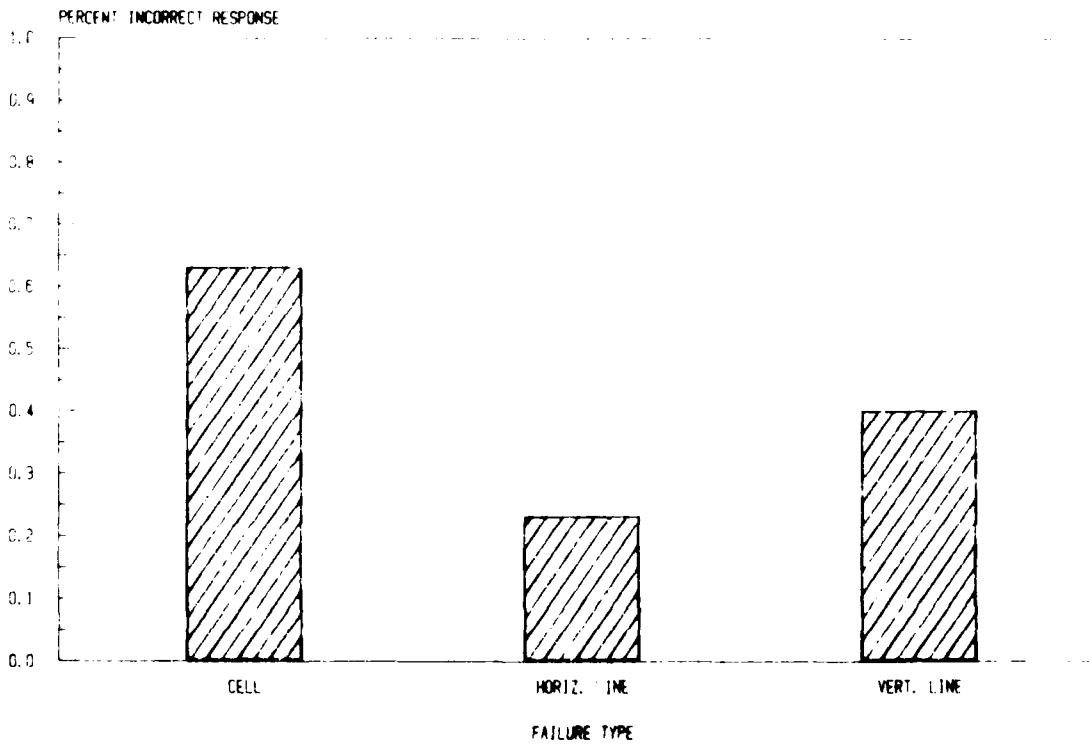
Figure 20. Effect of Failure Type on Incorrect Responses.
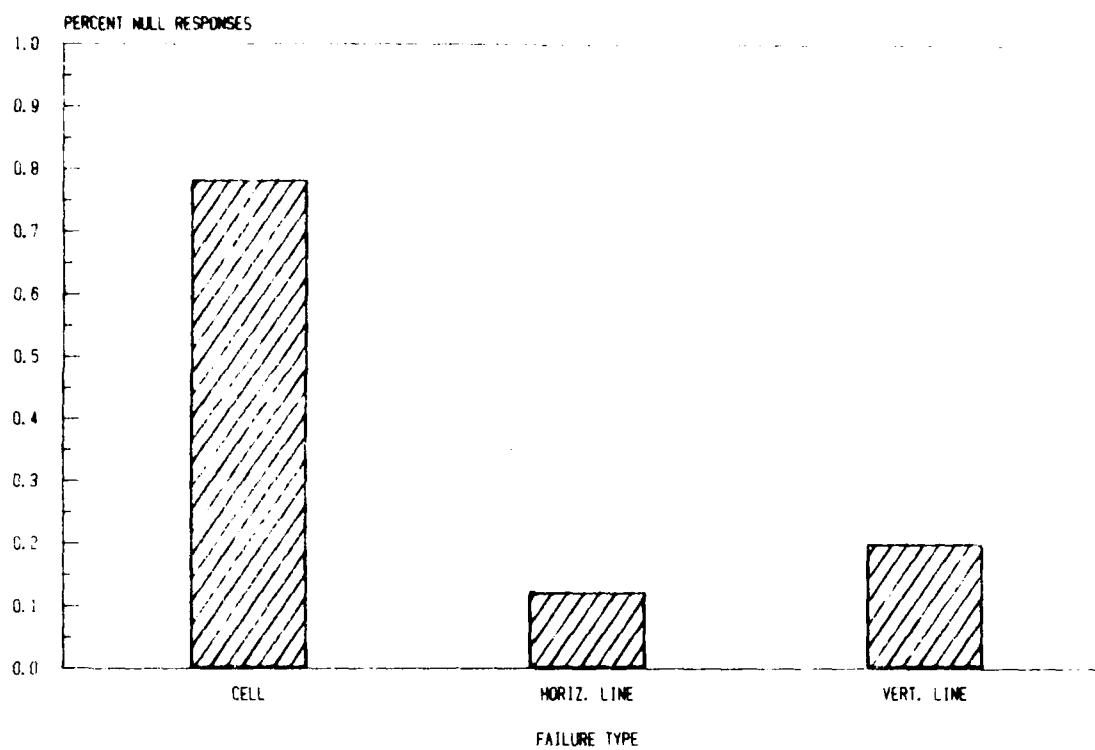
Figure 21.   Effect of Failure Type on Null Responses.
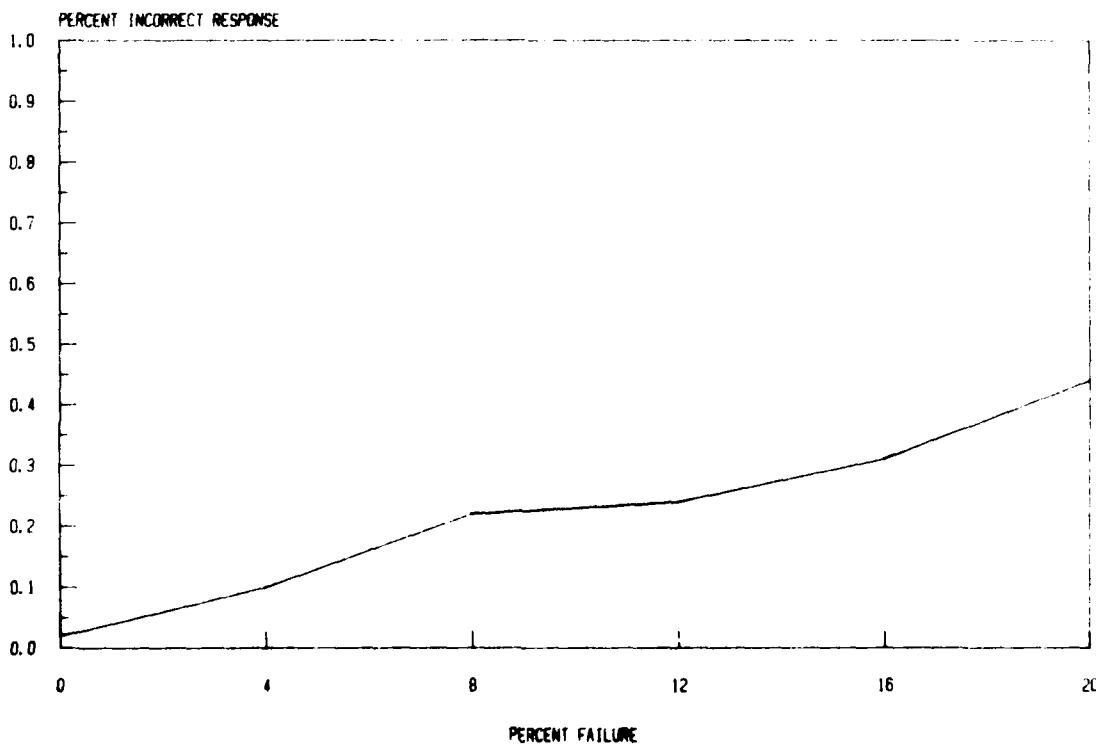
Figure 22.   Effect of Percent Failure on Incorrect Responses.

34

PERCENT NULL RESPONSES
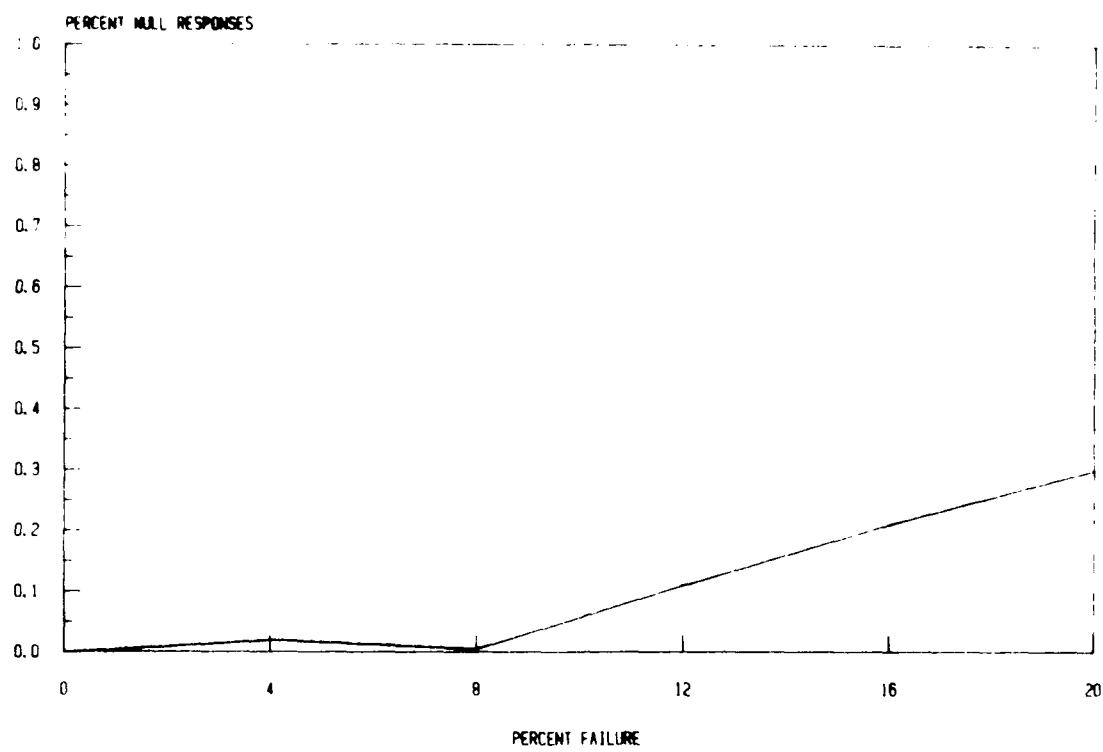
PERCENT FAILURE

Figure 23. Effect of Percent Failure on Null Responses.

PERCENT INCORRECT RESPONSE
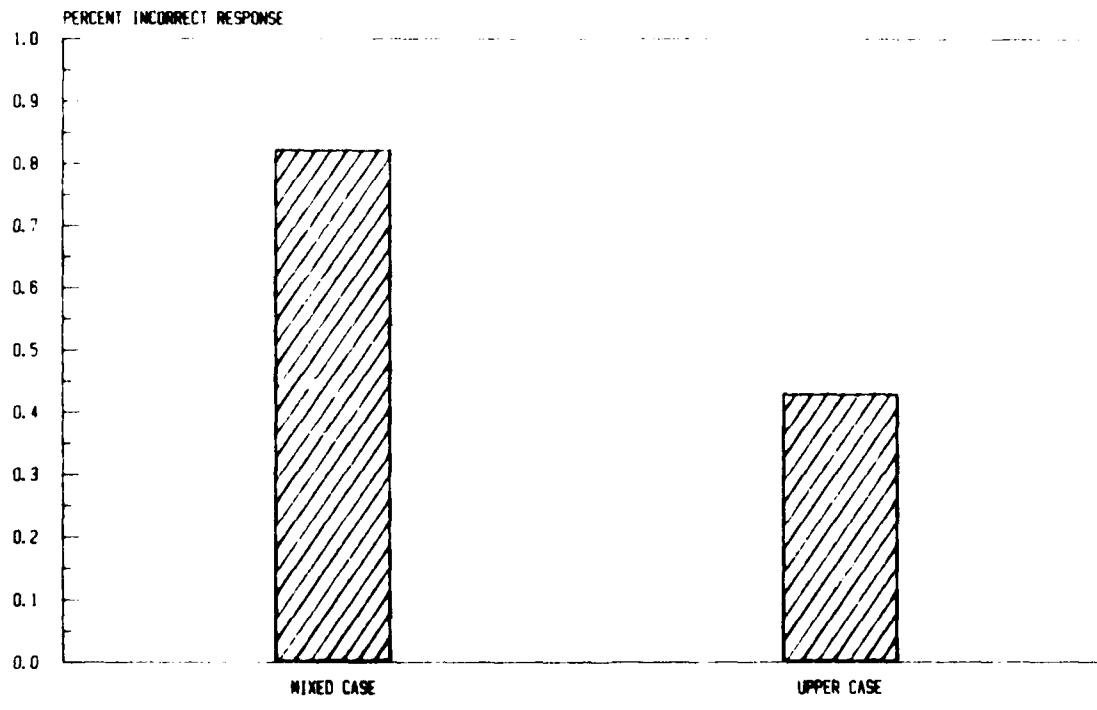
CASE

Figure 24. Effect of Case on Incorrect Responses.

PERCENT NULL RESPONSES

```
1.0 ┌─────────────────────────────────────────
    │
0.9 ┤      ┌──────┐
    │      │//////│
0.8 ┤      │//////│
    │      │//////│
0.7 ┤      │//////│
    │      │//////│
0.6 ┤      │//////│
    │      │//////│
0.5 ┤      │//////│
    │      │//////│
0.4 ┤      │//////│
    │      │//////│
0.3 ┤      │//////│                    ┌──────┐
    │      │//////│                    │//////│
0.2 ┤      │//////│                    │//////│
    │      │//////│                    │//////│
0.1 ┤      │//////│                    │//////│
    │      │//////│                    │//////│
0.0 └──────┴──────┴────────────────────┴──────┴──
            MIXED CASE                  UPPER CASE
```
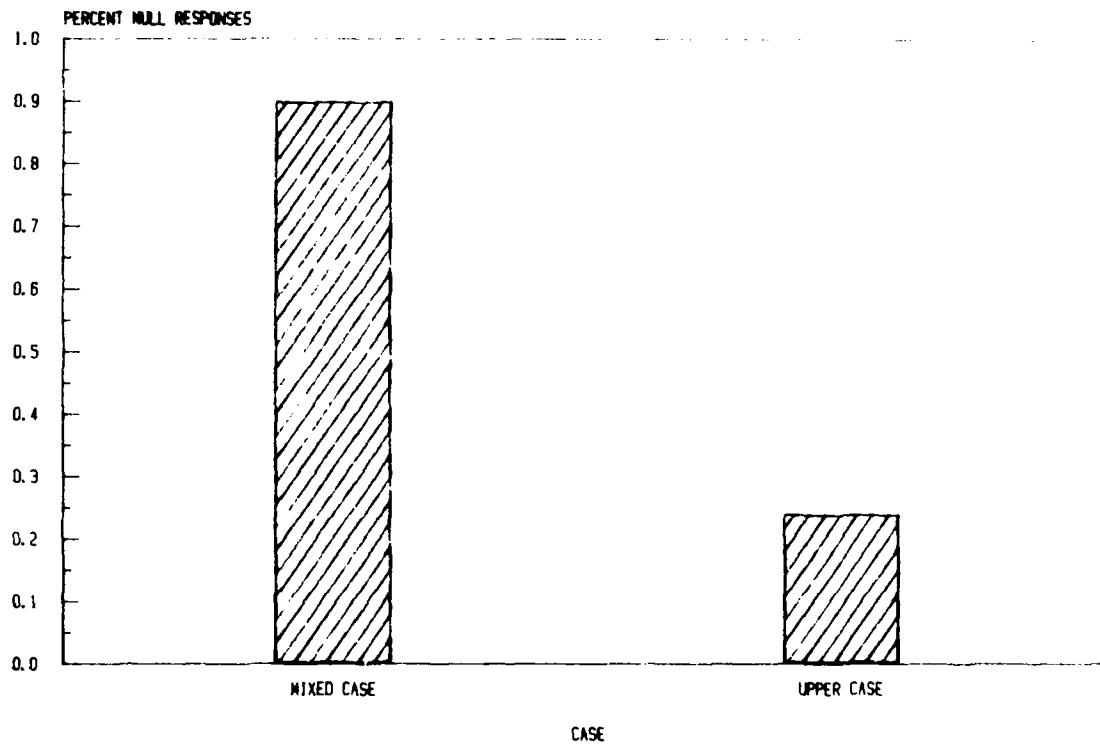
CASE

Figure 25.   Effect of Case on Null Responses.

DISCUSSION

Although the statistical analysis of the results of this factorial experiment is somewhat complex in its presentation, it is clear that the pertinent variables interact significantly to define their influence on display readability. Thus, interpretation of the influence of the variables must be made in the cautious awareness that such interactions are intrinsic, meaningful, and important in the determination of acceptable design levels.

Failure Mode

The on cell condition resulted in longer response times and more null responses than did the off cell condition. As random cells, horizontal lines, and vertical lines are increasingly added to or removed from the dot matrix characters, response times to read the passages increase. Additionally, while subjects did not differ in their incorrect response proportions for on vs. off conditions, they did make more null responses to the on condition. That is, incorrectly added cells or lines result in longer response times and more occasions on which the inappropriate word in the passage simply cannot be read at all.

Riley and Barbato (1978) found that character identification was not differentially affected by the removal of cells, addition of cells, or the simultaneous addition and removal of cells. They sought to determine experimentally the importance of each addressable pixel in a 7 X 9 matrix, containing 5 X 7 characters, and used these "importance" values to determine which cells should fail "on" and "off." They degraded one character at a time. The characters in this present experiment, on the other hand, were all randomly degraded in accordance with the experimental variables.

It is likely that the Riley and Barbato (1978) results differ from the present results because they degraded only one character at a time. Semple et al. (1971) suggested that since single characters do not generate any contextual cues, the findings for single symbols should not be generalized to words and vice versa. The legibility and readability of alphanumeric characters are affected by the informational content of alphanumerics when presented in the form of character strings or words, rather than as isolated or unrelated characters (Snyder and Taylor, 1979). The readability of a passage becomes a more difficult task when an entire word is missing than when only one letter of a word is missing.

It should also be noted that the effects of on and off failures are not the converse of one another. In the on condition, it is possible to add pixels to the entire display, except where pixels are already turned on for characters in the passage. In the off condition, it is possible to remove only pixels that are part of the characters in the passage, since the remaining pixels are already turned off. (In this experiment, sampling and determination of the "failed" pixels to be inappropriately turned on or off was done without regard to whether

they were already on or off.) Since the characters define fewer than half of the pixels in the passage area, the possible number of **displayed** failed cells is greater with normally off pixels turned on than with normally on pixels turned off. Thus, the differences between the on and off experimental conditions may well be due to the displayed failure frequencies, rather than to the underlying failed cell frequencies. Of course, this is a meaningful "real world" result and *not an experimental artifact*.

Failure Type

Randomly failed individual cells resulted in the poorest performance. Individual cells, whether on or off, were the most distracting failure conditions, producing the longest reading times and the most null and incorrect responses. Conversely, horizontal line failures resulted in the shortest response times and the smallest proportion of null and incorrect responses.

The blank areas on the display where no characters are written, i.e., between the lines of text, are larger than are the vertical areas between characters. Thus, random placement of horizontal lines will occur more frequently in non-text areas than will random placement of vertical lines. Furthermore, a subjective analysis of alphabetical characters suggests that there are more vertical line segments than there are horizontal line segments. Thus, horizontal line errors are less likely to coincide with horizontal character segments than are vertical line errors to coincide with vertical character segments.

Cell failures added to the display as on failures led to greater response times than did on line failures. However, cell off failures resulted in better performance than did line off failures. A subjective evaluation of these different display appearances indicates that the "salt and pepper" appearance of on cell errors are particularly difficult to "see through," whereas the missing cells in the off condition are not nearly as detectable. This difference between on and off conditions is not as apparent for the line errors.

Percent Failure

Throughout the entire experiment, the greater the percent failure the longer is the response time. Beyond the 8% failure level, both incorrect and null responses increase consistently. These results are quite consistent with those of Pastor and Uphaus (1982), who found significant reading failures in dot matrix numerics with only 2% dot loss.

Case

For printed text, reading speed is generally faster with combinations of upper and lower case. Furthermore, the mixed case presentation is generally found to be both more legible and more pleasing (Tinker, 1965). In this experiment, subjects took longer to read mixed case characters than upper case characters under all percent failure levels from 8% to 20%. At the error-free (0%) and 4% levels, the differences

were not significant (Table 4). Thus, the results do not contradict those from printed text research; however, the suggestion exists in these data that even failure-free passages may be read more slowly with mixed case than with upper case presentation (Figure 13). This effect, if supported by additional data, would not be surprising inasmuch as lower case dot matrix characters have relatively few dots with which to define each character, on the order of 7 X 7 with a 7 X 9 upper case matrix format. Thus, since it is well established that there exists a significant difference in legibility of 5 X 7 as compared to 7 X 9 upper case characters, one should expect a similar legibility difference between lower case 7 X 7 characters and upper case 7 X 9 characters.

Font

Although the font effect is not overall significant, the interactions involving the font variable are most enlightening. The Huddleston font yielded the best performance across all failure types and at greater percent failure levels. Thus, the Huddleston seems to be the most resistant to degradation. Coupled with other experimental results that show the Huddleston font to be at least as good as other dot matrix fonts, it appears that these data further support the recommendation that the Huddleston font be selected to maximize legibility and readability.

The only exception to this recommendation lies in the situation in which reversed polarity displays are used. For example, cathode-ray tubes designed to produce good image quality with positive contrast typically yield poorer image quality with negative contrast characters. The reason for this is that the visual system requires, for equal legibility, a broader strokewidth for dark characters on a light background (negative contrast) than it does for light characters on a dark background. Since most cathode-ray tube spot sizes are optimized for light characters, they produce thin, low-appearing-contrast dark characters on a light background. However, when the Huddleston font is used in this negative contrast mode, the double width diagonal lines in the Huddleston font (e.g., letters M, N, W) appear much darker than do the vertical and horizontal single width lines. This astigma⁻ism results in distractingly heavy diagonal lines in some Huddleston font letters and should be avoided for that reason. Of course, with the constant-pixel-width, light characters used in this study on the plasma display, this astigmatism did not exist.

SUMMARY AND DESIGN RECOMMENDATIONS

Flat panel displays frequently suffer from discrete failure conditions. The data from this experiment apply to operator performance under a variety of controlled failure conditions typical of flat panel displays. Therefore, the results can be used to estimate the conditions under which display readability will be degraded and the display should therefore be evaluated for replacement. Similarly, the data can be used to set quality assurance levels for production of flat panel displays or for acceptance sampling plans.

Based upon these data, the following recommendations are made for display acceptance and information presentation for situations where failures of display elements are likely.

Failures

The effect of display failures depends on the failure mode as well as on the type of failure. While line failures are more harmful than cell failures for off failed cells, cell failures are more harmful for on failed cells. In both cases, however, maintaining the percent of failed cells below 2% will not result in any appreciable degradation of readability. In actuality, this is a fairly noticeable level of degradation on a sizeable display and one which is unlikely to be acceptable for purely cosmetic reasons in spite of the fact that no performance degradation is noted.

Case

The upper case presentation resulted in significantly better performance, particularly under degraded display conditions. Thus, for dot matrix displays, all upper case is recommended except for situations such as text editing in which mixed case is imperative.

Font

The Huddleston font is most resistant to degradation and has as high a legibility as any other dot matrix font investigated to date. Unless a negative contrast display with significant astigmatism due to a thin strokewidth is used, the Huddleston font is recommended.

REFERENCES

Albert, D. E. Prediction of intelligibility of contextual and noncontextual dot matrix characters. Unpublished M. S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1975.

Bauer, H. J. Some solutions of visibility and legibility problems in changing speed signs. Washington: National Academy of Sciences, National Research Council, Highway Research Board Bulletin, 1962.

Burnette, J. T. Optimal element size-shape-spacing combinations for a 5 X 7 dot matrix visual display. Unpublished M. S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1976.

Carver, R. P. and Tinker, M. A. **Basic reading rate scale.** Minneapolis: Revrac, 1970.

Cornog, D. Y. and Rose, F. C. Legibility of alphanumeric characters and other symbols: II. A reference handbook. Washington: U. S. Government Printing Office, National Bureau of Standards Miscellaneous Publication 262-2, 1967.

Huddleston, H. F. An initial examination of 5 x 7 matrix alphanumerics for digital television. Institute of Aviation Medicine, Royal Air force, IAM Report No. 496, 1970.

Long, E. R., Reid, L. S., and Queal, R. W. Factors determining the legibility of letters and words derived from elemental printers. Wright-Patterson Air Force Base, Ohio: Wright Air Development Center Technical Report 5924, 1951.

Maddox, M. E., Burnette, J. T., and Gutmann, J. C. Font comparisons for 5 x 7 dot matrix characters. **Human Factors,** 1977, **19,** 89-93.

McCormick, E. J. and Sanders, M. S. **Human factors in engineering and design.** New York: McGraw-Hill, 1982.

Pastor, J. R. and Uphaus, J. S. Significant reading failures in 7 x 9 dot matrix ASCII numbers with two percent dot loss. **Society for Information Display Digest,** 1982, 198-199.

Riley, T. M. and Barbato, G. J. Dot-matrix alphanumerics viewed under discrete element degradation. **Human Factors,** 1978, **20,** 473-480.

Scanlan, L. A. and Carel, W. L. Human performance evaluation of matrix displays. Wright-Patterson Air Force Base, Ohio: Technical Report AMRL-TR-76-39, 1976.

Semple, C. A., Jr., Heapy, R. J., Conway, E. J., Jr., and Burnette, K. T. Analysis of human factors data for electronic flight display systems.

Wright-Patterson Air Force Base, Ohio: Final Technical Report, Contract F33615-70-1132, 1971.

Shurtleff, D. A. Studies of display legibility. XII: The relative legibility of symbols formed from matrices of dots. Air Force Systems Command, Electronic Systems Division Technical Report ESD-TR-69-432, 1970.

Spencer, H., Reynolds, L., and Coe, B. The effects of image degradation and background noise on the legibility of text and numerals in four different typefaces. Readability of Print Research Unit, Royal College of Art, 1977.

Snyder, H. L. Human visual performance and flat panel display image quality. Blacksburg, VA: Virginia Polytechnic Institute and State University Technical Report HFL-80-1, 1980.

Snyder, H. L. and Maddox, M. E. Information transfer from computer-generated dot-matrix displays. Blacksburg, VA: Virginia Polytechnic Institute and State University Technical Report HFL-78-3, 1978.

Snyder, H. L. and Taylor, G. B. The sensitivity of response measures of alphanumeric legibility to variations in dot matrix display parameters. **Human Factors**, 1979, **21**, 457-471.

Tinker, M. A. **Bases for effective reading**. Minneapolis: Minnesota Press, 1965.

Vartabedian, A. G. The design of visual displays. Holmdel, N.J.: Bell Laboratories, 1970.

Vartabedian, A. G. Developing a graphic set for cathode ray tube display using a 7 x 9 dot pattern. **Applied Ergonomics**, 1973, **4**, 11-16.

APPENDIX A. EXAMPLES OF TINKER READING PASSAGES

Howard could always be found at the printer's shop in his spare time, for he loved to run errands for the staff. We knew he was a baker at heart.

Uncle Time gave Micky a new pair of roller-skates, and as she went down the street she called to the mailman, "See how fast I go on my new sled."

Suddenly a strong gust of wind blew Harry's hat off his head and dropped it in a mud puddle. Harry was annoyed for it was the best coat he owned.

Jean made some delicious muffins for her father's breakfast, and he was so pleased he said he would give her a dollar every time she made such good pictures.

APPENDIX B.   INSTRUCTIONS

The display located in front of you is a plasma panel. You may not be familiar with such a display but it has many capabilities similar to an ordinary CRT. You will notice a triggering device in front of the panel. Once the experiment has begun you may press the trigger and release it. A passage will appear on the plasma panel. If you press the trigger again and release it, the passage will disappear. Now try it. **GENTLY** press the trigger and release it.

Simple one or two sentence passages called Tinker reading passages will be displayed on the plasma panel. You have already seen one in the previous example. Here is another example.

The bright colors in Barbara's dress ran when her mother washed it with the rest of the clothing. Mrs. Atkins said she had never seen a boy run so much.

Again, notice that one work does not fit in the sentence. In this case the critical word is "boy." Your task in this experiment is to tell the experimenter aloud what the critical word is. Try to state the critical word as clearly as you can so that the experimenter will accurately understand you.

Seven practice trials will be given to you. You will notice that sometimes extra "dots" may appear in the passage, or that some "dots" may be missing. There are seven different ways extra dots may appear, may be missing, or may not be missing at all. Sometimes more dots may appear on a particular trial than on another trial, and sometimes more dots may be missing on one trial than on another trial. The steps in the experiment are:
1. You will see a rectangular box on the plasma panel display screen.
2. Once you see this box appear press the trigger.
3. Release the trigger.
4. A passage will be displayed.
5. Read the passage.
6. Press the trigger.
7. Release the trigger.
8. Tell the experimenter what the critical word is.
9. Repeat the sequence when the box reappears.

Please try to read the sentence as quickly and accurately as possible. When the plasma panel is turned on the power supply (white box on your left) will generate some noise. Just ignore this noise. The lights in the room will be turned off during the experiment. Before the practice session you will be given a few seconds to become accustomed to the dark.

Before we begin with the practice trials do you have any questions? Remember the seven steps. You may refer to them.

44

<<Practice Session>>

Are there any questions?   The experiment will take about one and a half
hours.   We will take a short rest period about half way through.   OK,
let's begin.

LMED
_8